# Evolutionary Optimisation of Real-Time Many-Cores

Leandro Soares Indrusiak

Real-Time Systems Group
Department of Computer Science
University of York
United Kingdom

# Introduction

- Manycore processor design and configuration requires decisions across a wide range of aspects
  - core selection and placement
  - DVFS settings
  - memory partitioning
  - task mapping
  - task priority assignment
  - security
  - …

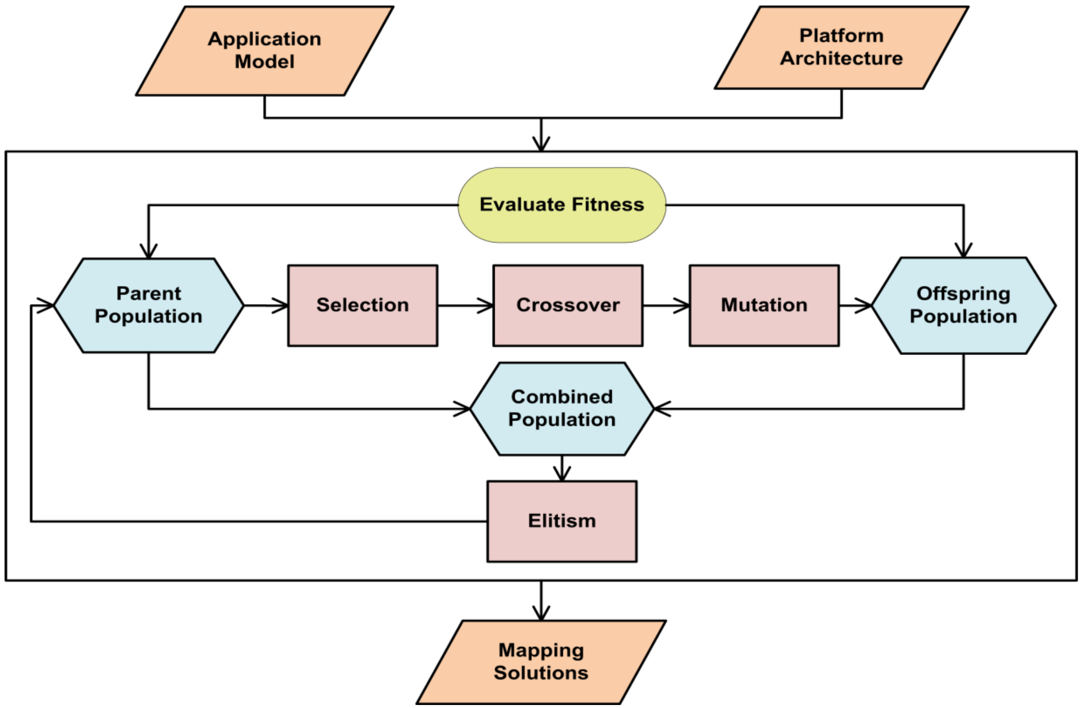
# Introduction

- Design and configuration for real-time manycores even harder
  - each decision may jeopardise real-time requirements
  - many trade-offs
  - optimisation problem is unlikely to be solved using only designer experience

- Use schedulability tests to guide the design space exploration
  - can be used as fitness function in a search-based evolutionary optimisation
  - guides the search towards full schedulability
  - much faster than simulation, therefore can cover a wider search space

# Optimisation Algorithm

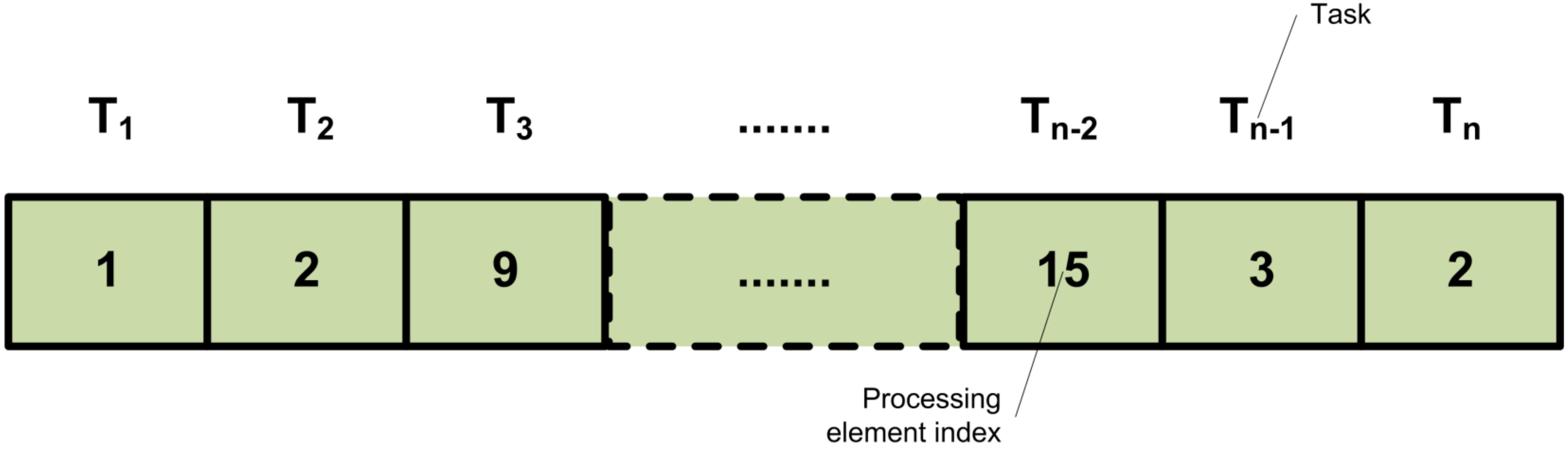
- Optimisation performed by a population-based evolutionary algorithm




P. Mesidis and L. S. Indrusiak, “Genetic mapping of hard real-time applications onto NoC-based MPSoCs — A first approach,” in Int Workshop on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC), 2011.
M. N. S. M. Sayuti and L. S. Indrusiak, “Real-time low-power task mapping in Networks-on-Chip,” in IEEE Computer Society Annual Symposium on VLSI (ISVLSI), 2013.

# Optimisation Algorithm

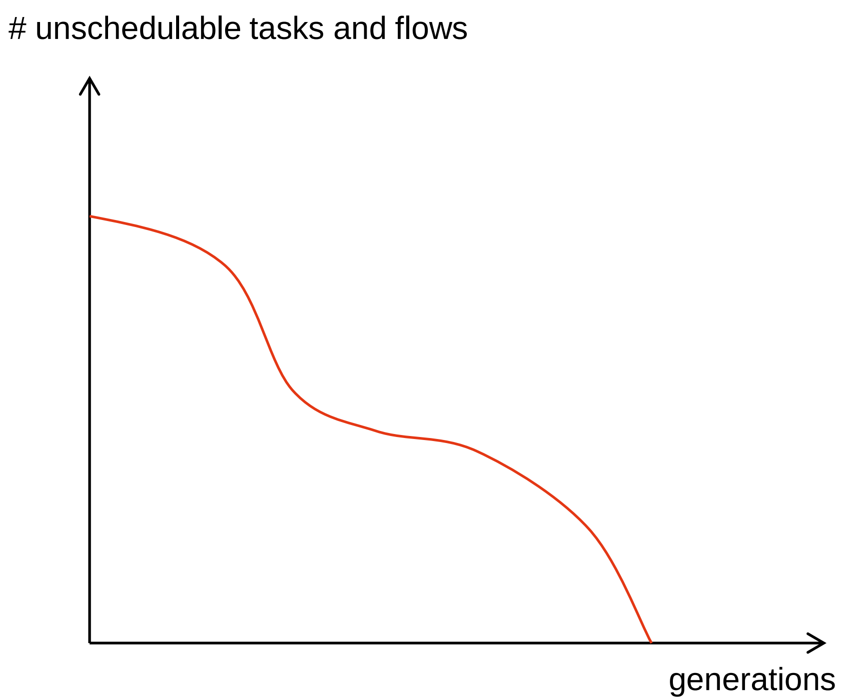
- A population contains a group of individuals represented by a chromosome structure




- Creation of new individuals is facilitated by operators:
  - Selection
  - Crossover
  - Mutation

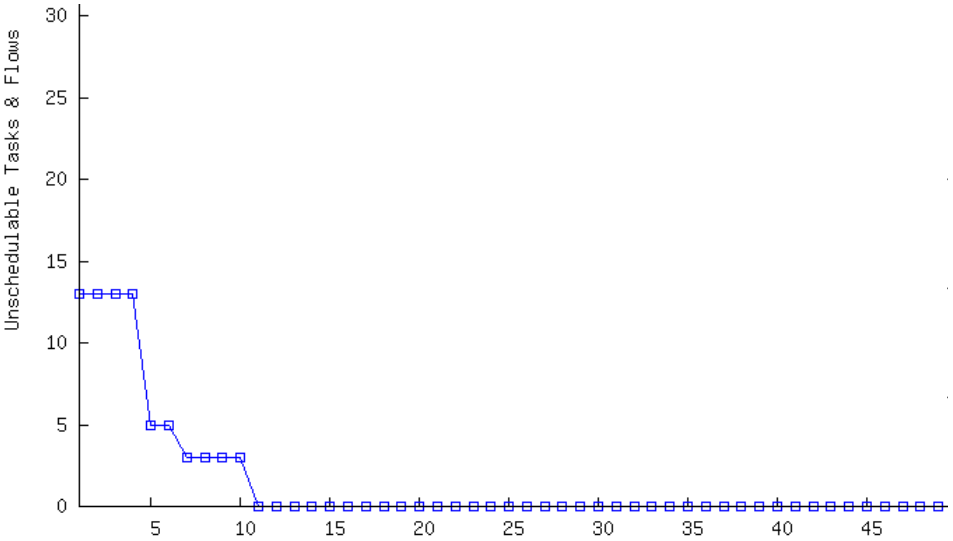
# Optimisation algorithm

- Goal: evolve a fully schedulable mapping over generations



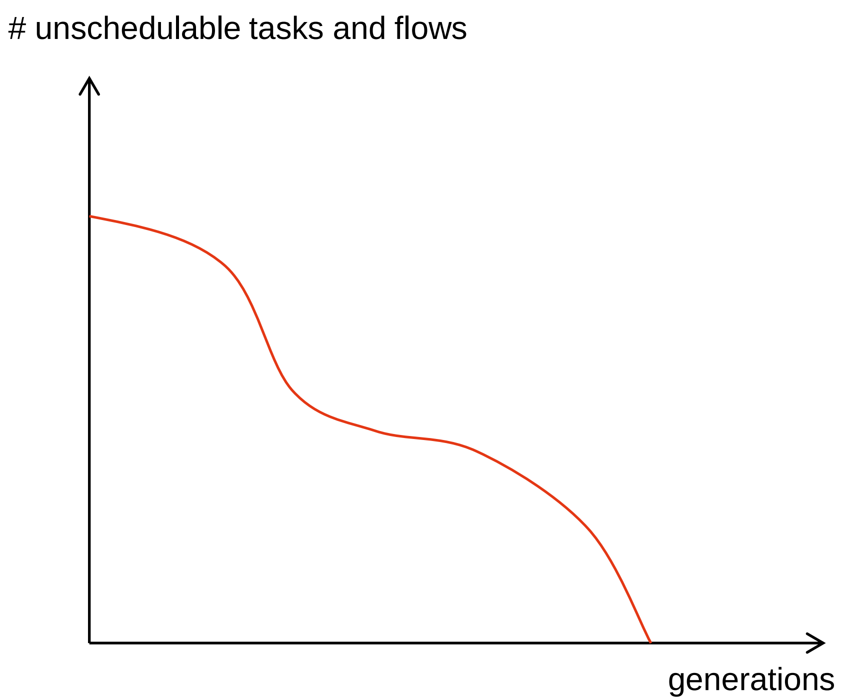
# Experiment results

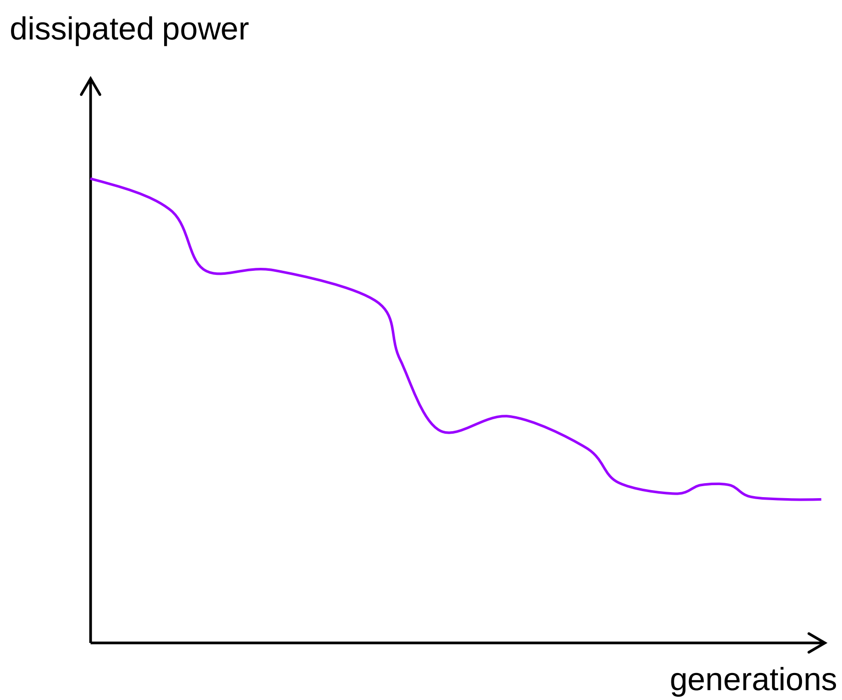
- Autonomous vehicle application (AVA) benchmark (38 communicating tasks), 4x4 Mesh NoC



# Multi-objective optimisation algorithm

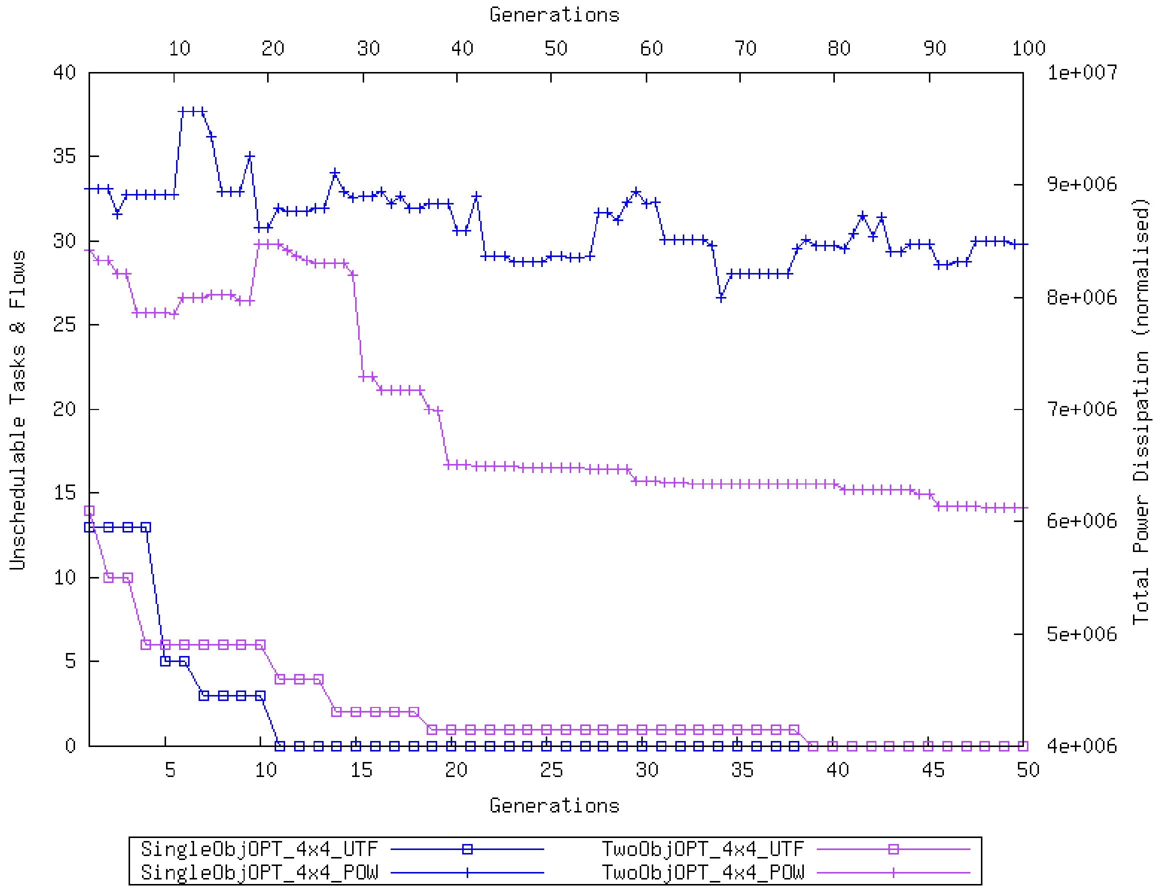
- Additional fitness functions can be added to the evolutionary algorithm
    - example: evolve mappings that are fully schedulable and low power



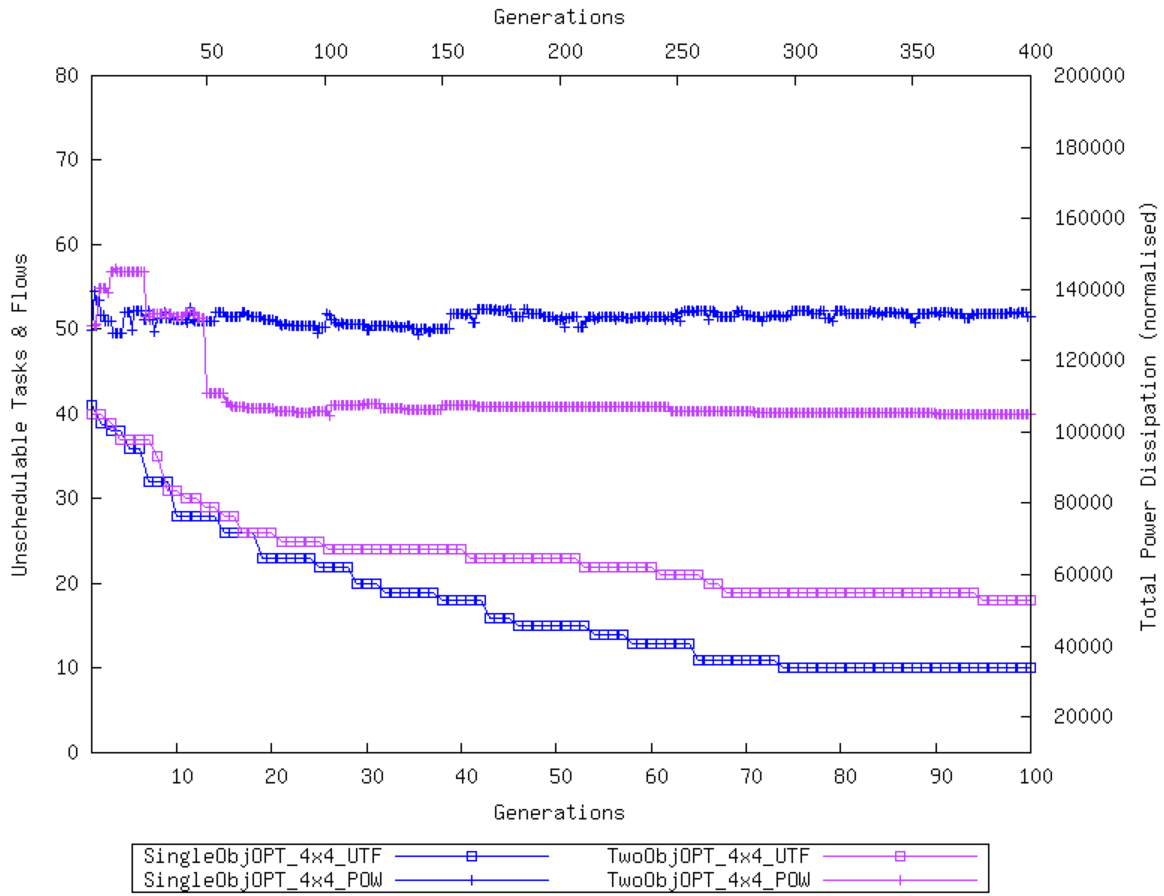
# Experiment results

- Autonomous vehicle application (AVA) benchmark, 4x4 Mesh
- Comparison of best solution convergence between single and multiple objectives



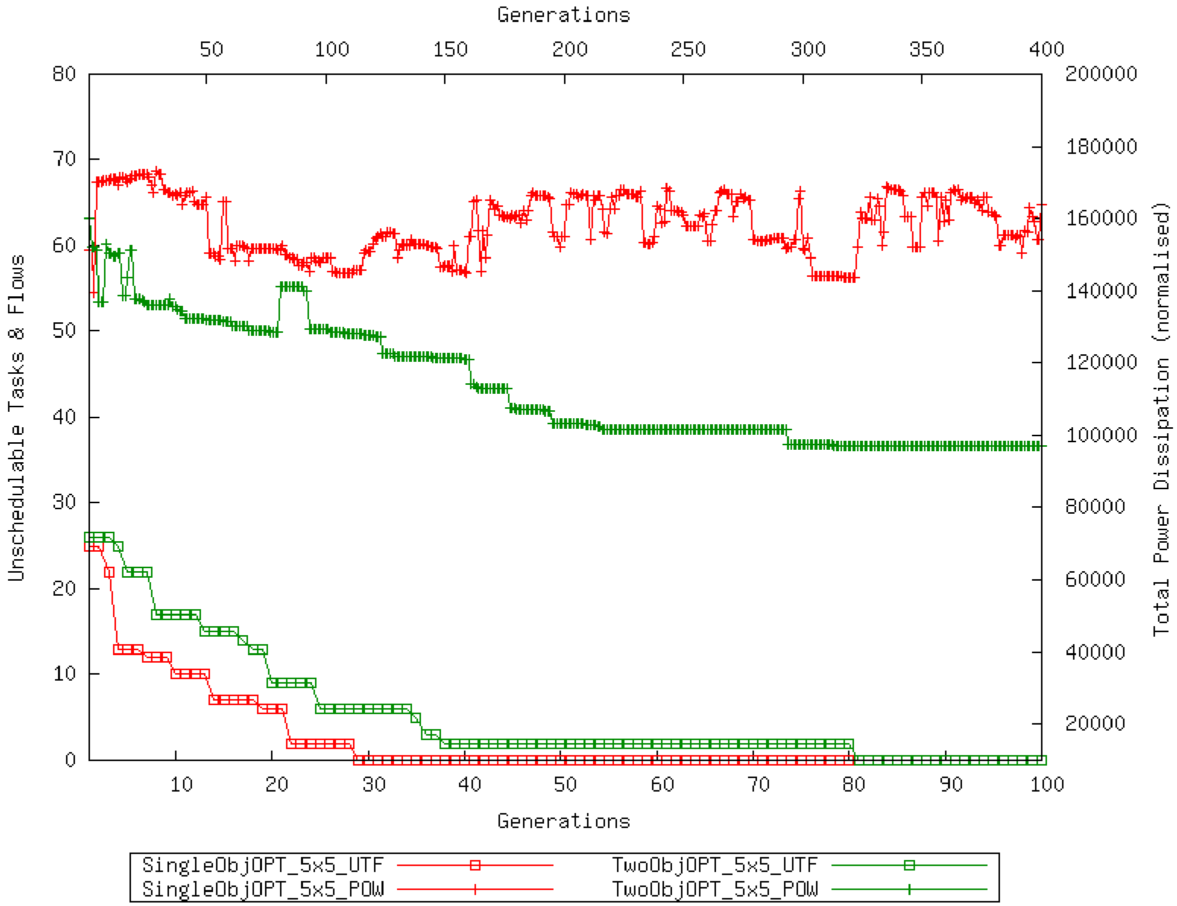
# Experiment results

- Synthetic application (SA) benchmark, 4x4 Mesh
- Comparison of best solution convergence between single and multiple objectives



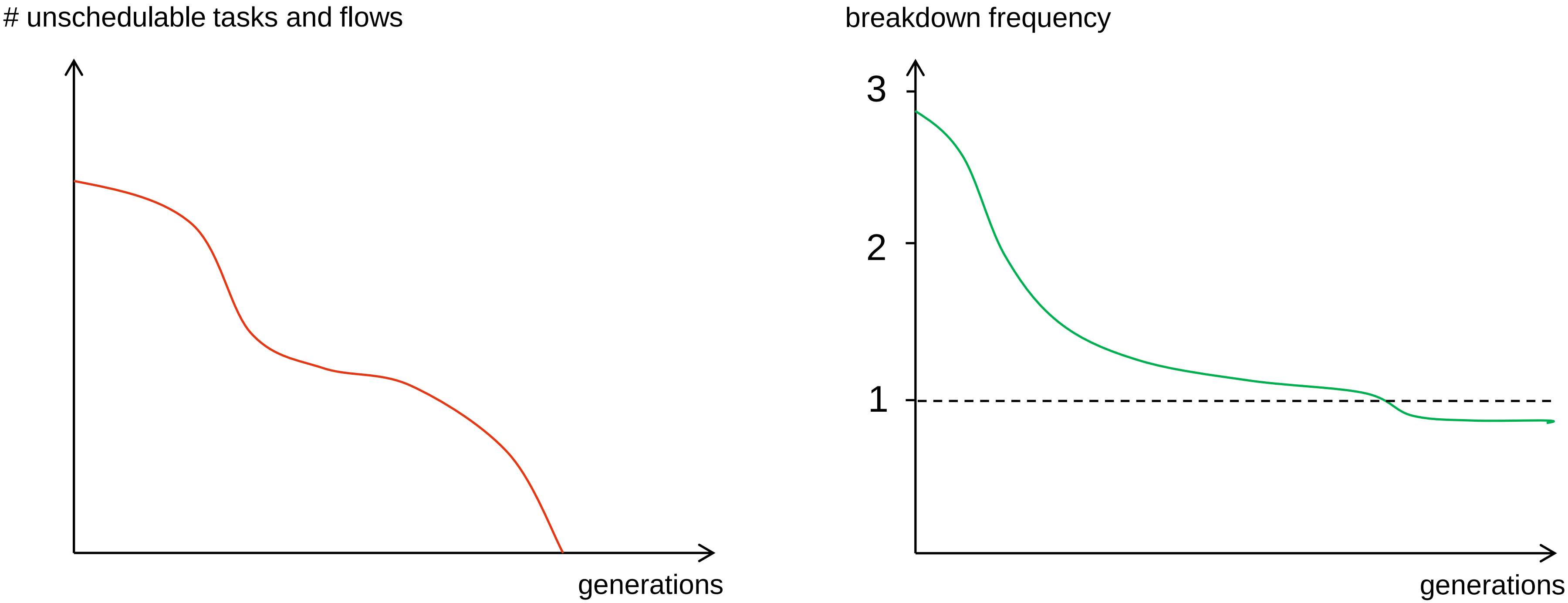
# Experiment results

- Synthetic application (SA) benchmark, 5x5 Mesh
- Comparison of best solution convergence between single and multiple objectives



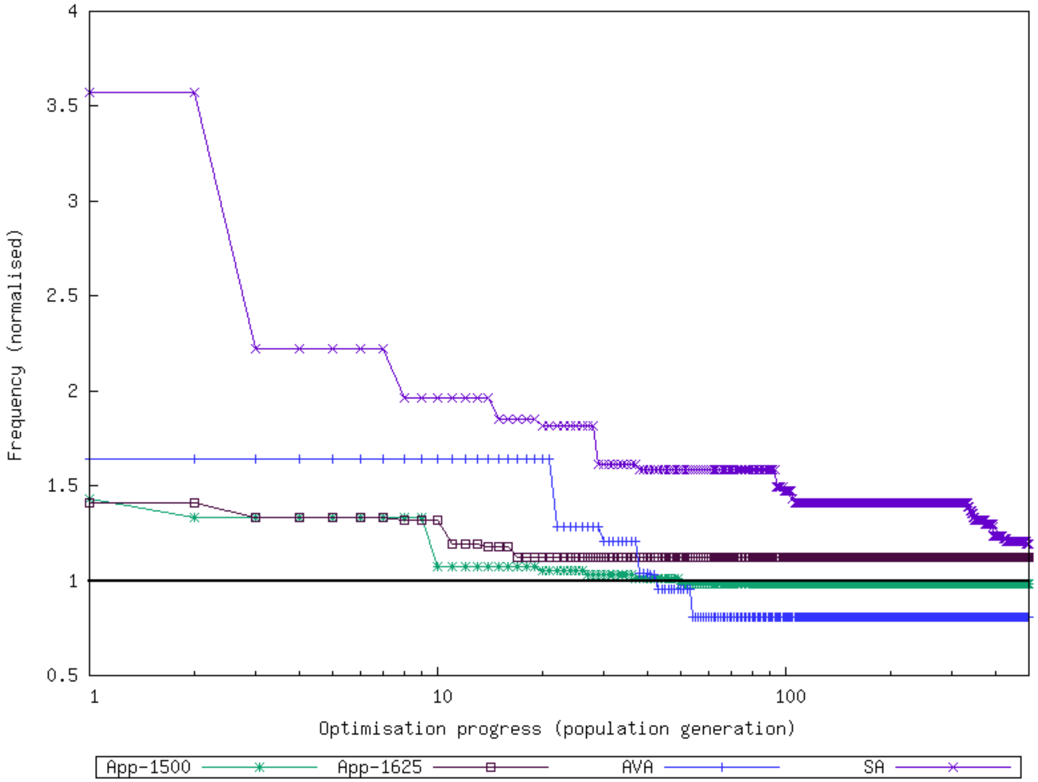
# Fitness Function

- An improved fitness function for schedulability based on the notion of breakdown frequency can be applied
  - evolve mappings with lower breakdown frequency



# Experimental Results

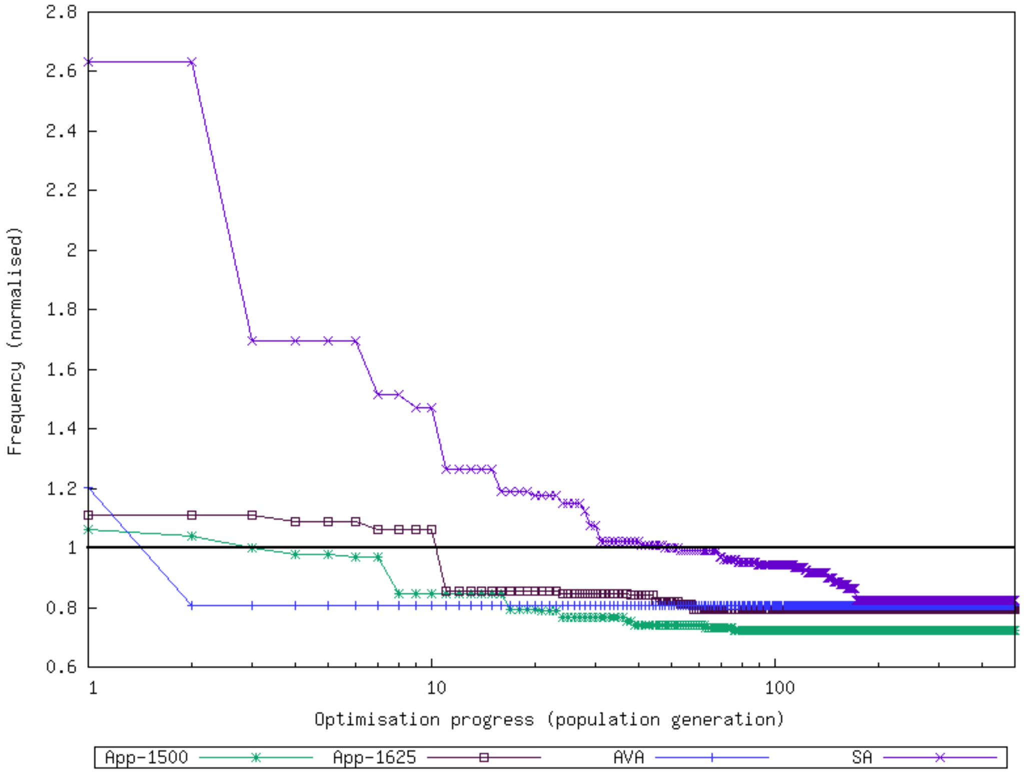
- AVA, SA, App-1500, App-1625 mapped onto 4x4 Mesh NoC
- Successful convergence of GA-BF




M. N. S. M. Sayuti and L. S. Indrusiak, "A Function for Hard Real-Time System Search-Based Task Mapping Optimisation," In 18th Int Symposium on Real-Time Distributed Computing (ISORC), 2015.

# Experimental Results

- AVA, SA, App-1500, App-1625 mapped onto 5x5 Mesh NoC
- Successful convergence of GA-BF

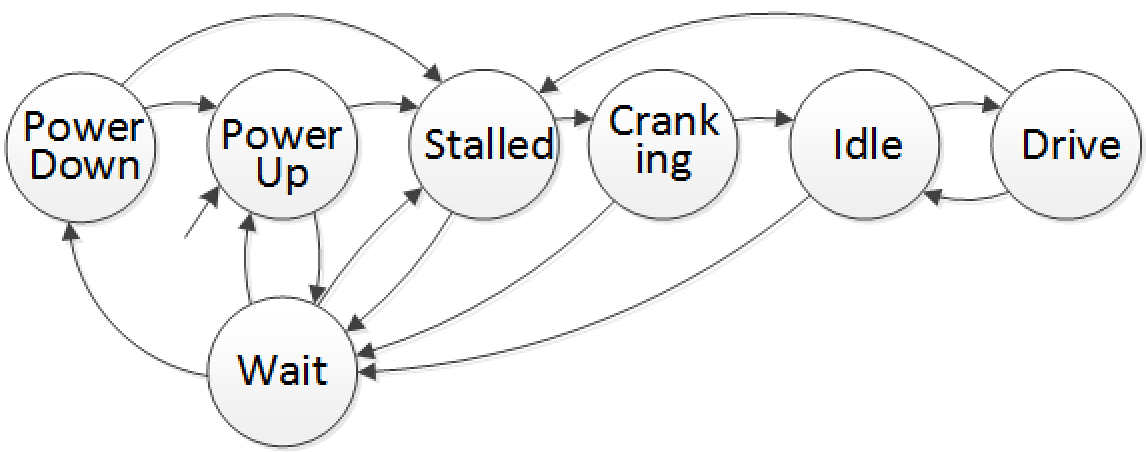



M. N. S. M. Sayuti and L. S. Indrusiak, "A Function for Hard Real-Time System Search-Based Task Mapping Optimisation," In 18th Int Symposium on Real-Time Distributed Computing (ISORC), 2015.

# Multi-mode applications

- Applications can have identifiable operation modes with different behaviours
  - tasks may only be active in some modes
  - task execution times or communication volumes may differ in different modes

- Example:

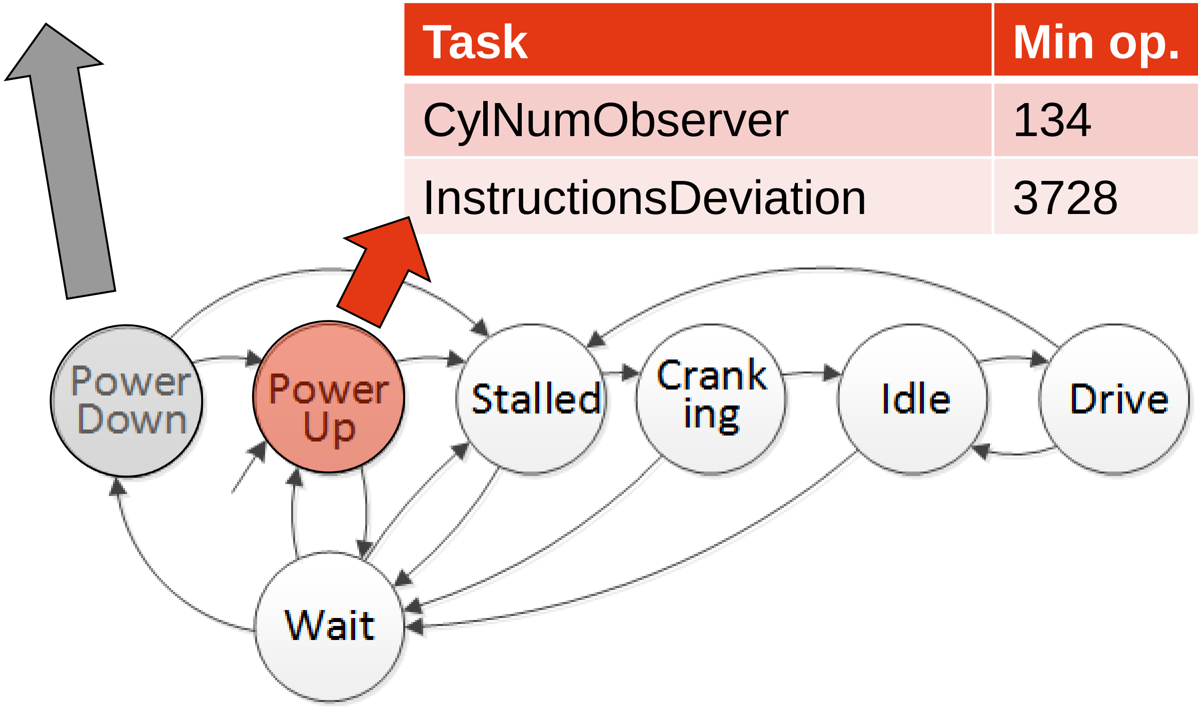



source: Bosch

# Multi-mode applications

| Task | Min op. | Max op. |
|---|---|---|
| CylNumObserver | 245 | 543 |
| InstructionsDeviation | 728 | 921 |

| Task | Min op. | Max op. |
|---|---|---|
| CylNumObserver | 134 | 345 |
| InstructionsDeviation | 3728 | 5921 |



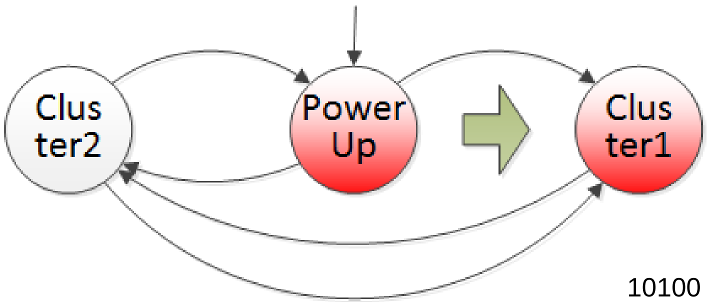

source: Bosch

# Multi-mode applications

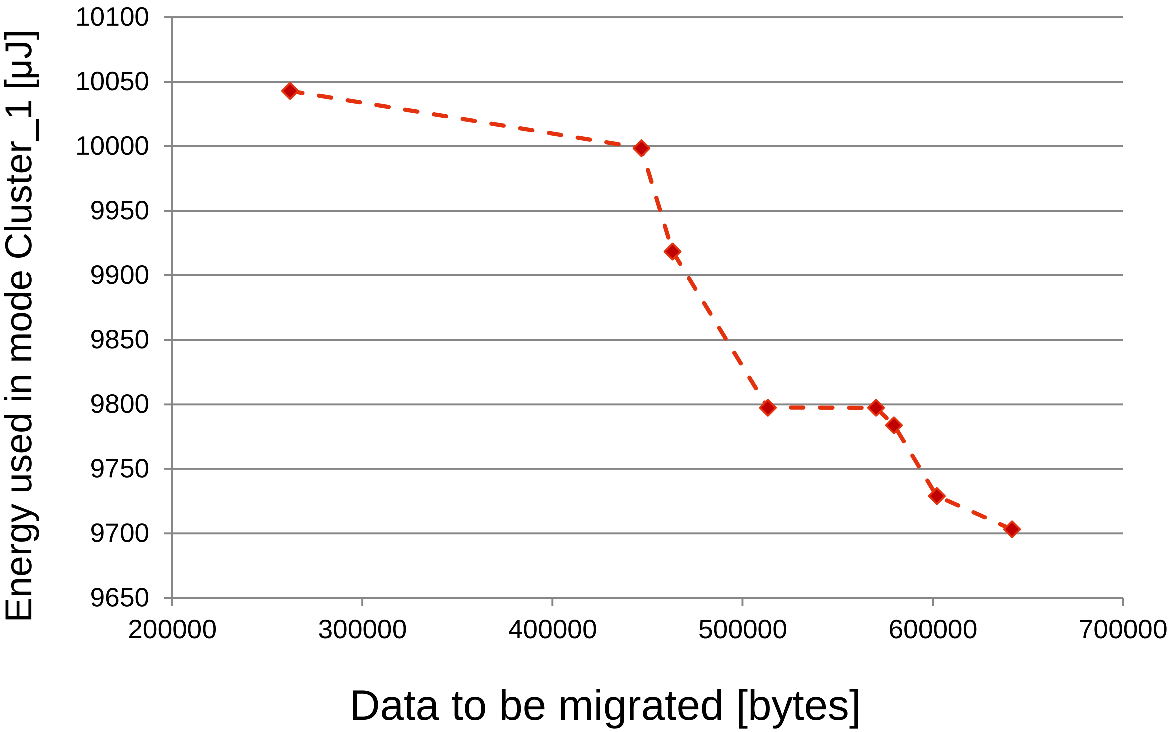
- Goal: evolve fully schedulable mappings for each mode, while
  - minimising task migration overheads during mode changes
  - keeping the system fully schedulable during mode changes
  - minimise energy dissipation
    - use as few cores as possible

P. Dziurzanski, A. K. Singh, L. S. Indrusiak, “Energy-Aware Resource Allocation in Multi-mode Automotive Applications with Hard Real-Time Constraints”, In 19th Int Symposium on Real-Time Distributed Computing (ISORC), 2016. (Best paper award)

# Experimental results – 2x2 NoC

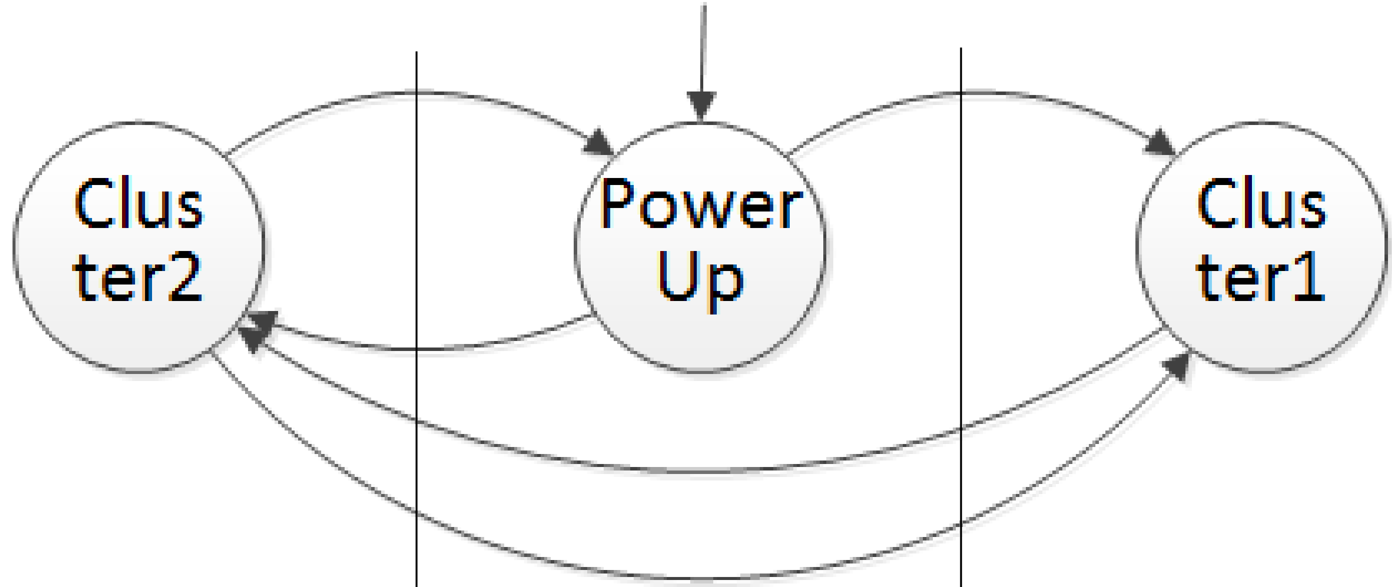



No deadline misses.




P. Dziurzanski, A. K. Singh, L. S. Indrusiak, "Energy-Aware Resource Allocation in Multi-mode Automotive Applications with Hard Real-Time Constraints", In 19th Int Symposium on Real-Time Distributed Computing (ISORC), 2016. (Best paper award)

# Experimental results – 2x2 NoC




| Cluster2 | Power Up | Cluster1 |
|---|---|---|
| • 2 idle cores<br>• 10 idle links | • 3 idle cores<br>• 16 idle links | • 0 idle cores<br>• 10 idle links |
| 5909.37 µJ per hyperperiod | 3093.01 µJ per hyperperiod | 9719.45 µJ per hyperperiod |

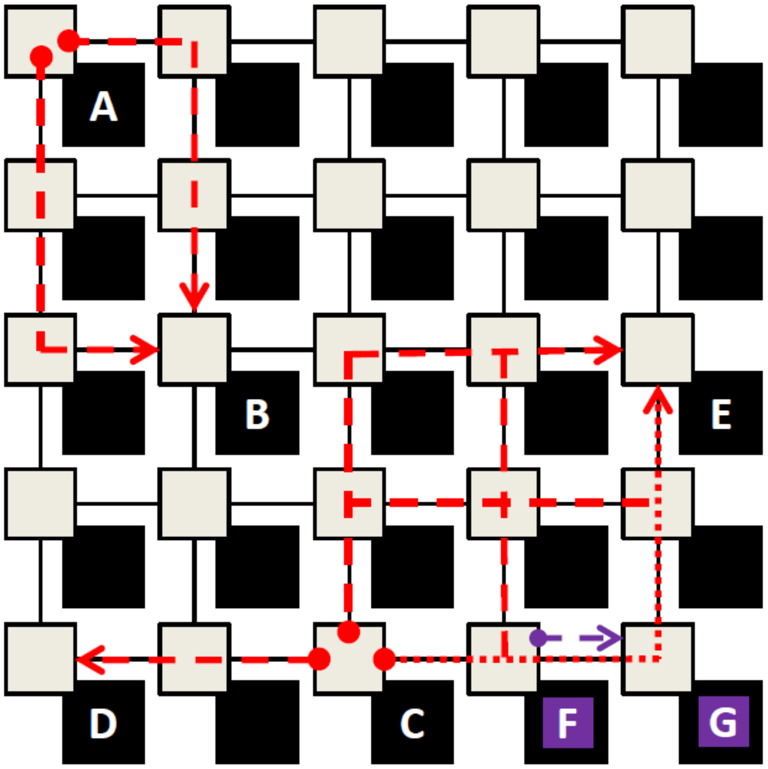
P. Dziurzanski, A. K. Singh, L. S. Indrusiak, “Energy-Aware Resource Allocation in Multi-mode Automotive Applications with Hard Real-Time Constraints”, In 19th Int Symposium on Real-Time Distributed Computing (ISORC), 2016. (Best paper award)

# Resilience against security attacks

- Side channel attacks can be used to obtain critical information in secure systems
  - timing attacks, e.g. task execution times or packet latencies can provide information that can significantly reduce the search space for brute-force attacks against private-key cryptography
  - energy/temperature side channels
- Goal: Reduce correlation between inter-process communication and packet latencies/energy footprint, while respecting communication deadlines

# Resilience against security attacks

- Approach: packet route randomisation at runtime
  - routers select route from a predefined set of choices
  - semi-adaptive routing algorithms such as west-first can still guarantee deadlock freeness
  - latency variabilty can increase significantly
  - energy hotspots become less predictable



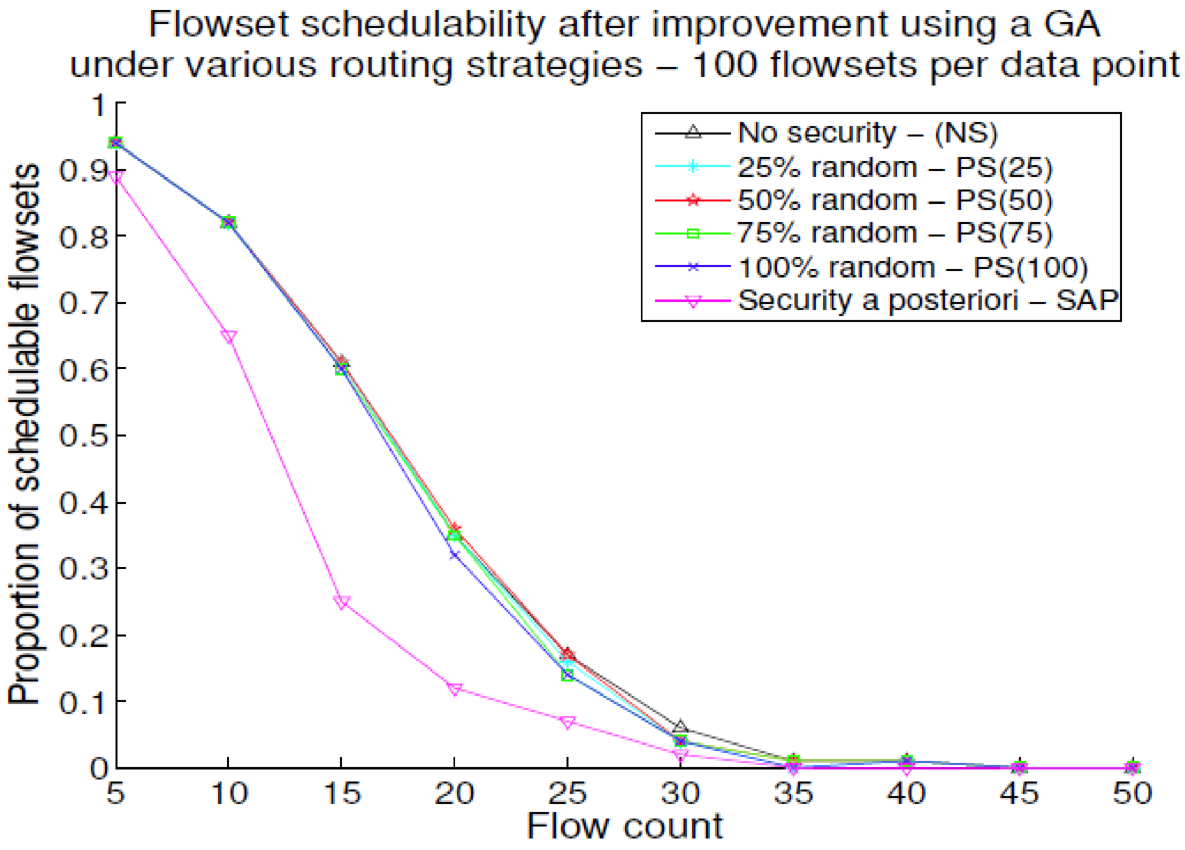
# Resilience against security attacks

- Problem: latencies should never exceed deadline
  - solution: extend schedulability test so that it considers interferences over all possible randomly-selected routes
    - not trivial, but simplifications are possible
    - use extended schedulability test to guide a search for a configuration where all packets meet their deadlines even under the additional interferences created by route randomisation
    - search could have an additional objective that aims to maximise randomisation as long as schedulability is not sacrificed

L. S. Indrusiak, J. Harbin, M. J. Sepúlveda, "Side-channel attack resilience through route randomisation in secure real-time Networks-on-Chip", In Int Symposium on Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC), 2017. (Best paper award)

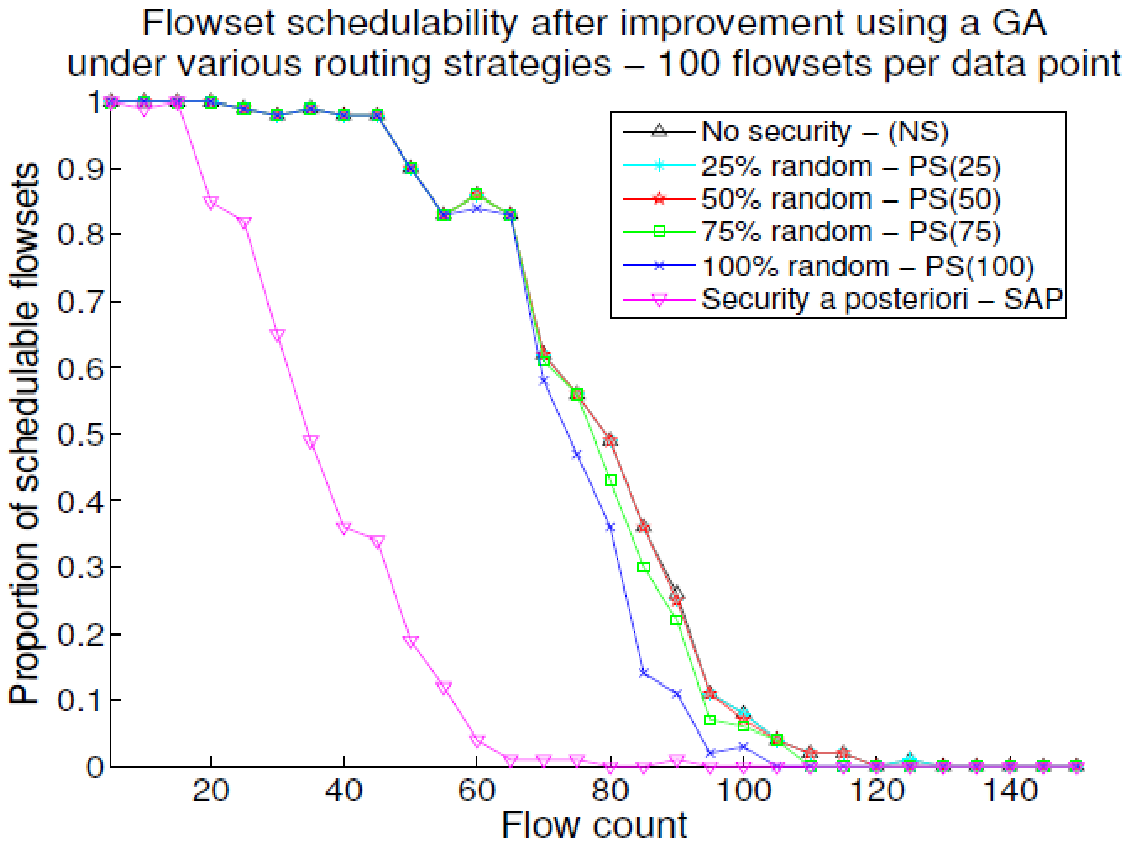
# Experimental results

- 4x4 NoC



# Experimental results

- 8x8 NoC



# Conclusions

- Schedulability tests are useful guides to evolutionary search heuristics attempting to configure multiple system aspects

- Published approaches addressing
  - task mapping for time (ReCoSoC 2011, 2012), energy (ISVLSI 2013, RTNS 2013), multi-mode operation (RTNS 2015, ISORC 2016, EURASIP JES 2017)
  - priority assignment (PDP 2015)
  - packet routing (RTCSA 2016)
  - security (ReCoSoC 2017)
  - memory (PDP 2018)
  - hardware-accelerated evolutionary pipeline (GECCO 2016)