



# FP7 STREP EUROCLOUD: Reliability Implications on TCO of a Datacenter

**Isidoros Sideris**, Yiannakis Sazeides, Nikolas Ladas,  
Costas Kourougiannis  
University of Cyprus

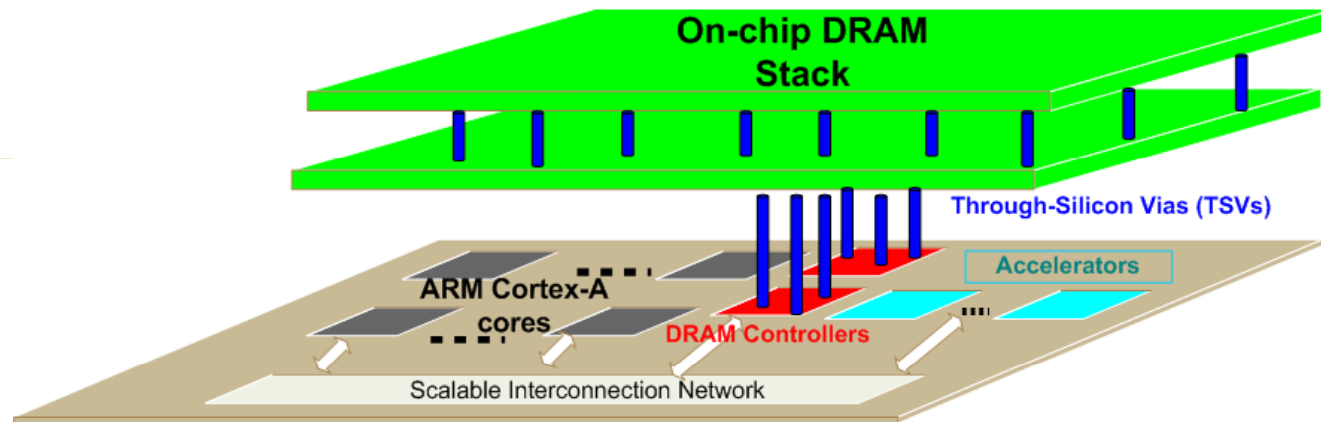
Veerle Desmet  
Ghent University

October 20, 2010  
Barcelona, Spain



# FP7 EuroCloud Project: Energy-conscious 3D Server-on-Chip for Green Cloud Services

## 3D Server-on-Chip



- FP7 STREP: 2010-2012
- Partners: ARM, Nokia, IMEC, UCY, EPFL
- Coordinator: ARM
- [www.eurocloudserver.com](http://www.eurocloudserver.com)
- UCY leads Reliability Work Package

# Analytical Framework for Availability/TCO estimation

- Motivation
  - Build a tool that can guide us in Reliability Design Considerations of the Server-on-Chip.
  - Evaluate different solutions, analytically, without long microarchitectural simulations or monte carlo experiments.
  - Minimize the objective function (=TCO)
- Inputs:
  - Architecture of the Server-on-Chip
  - Workload requirements
  - Accesses of various structures (which are the hot entries of an array?)
  - Characterization of circuit components (e.g. pfail of SRAM cells for a certain voltage)
  - Other parameters (MTTF for power supplies, board, PCB etc)
  - Reliability configuration employed
- Output:
  - TCO for a certain availability
  - Enviromental impact

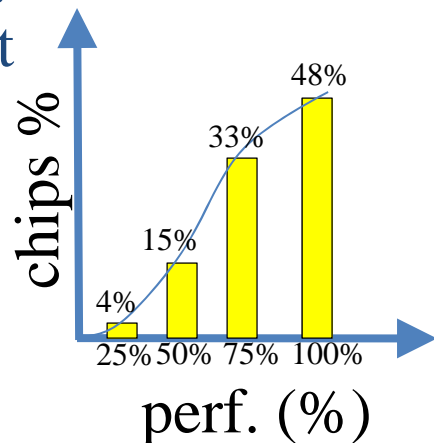


# Implications of Reliability on Performance/Availability

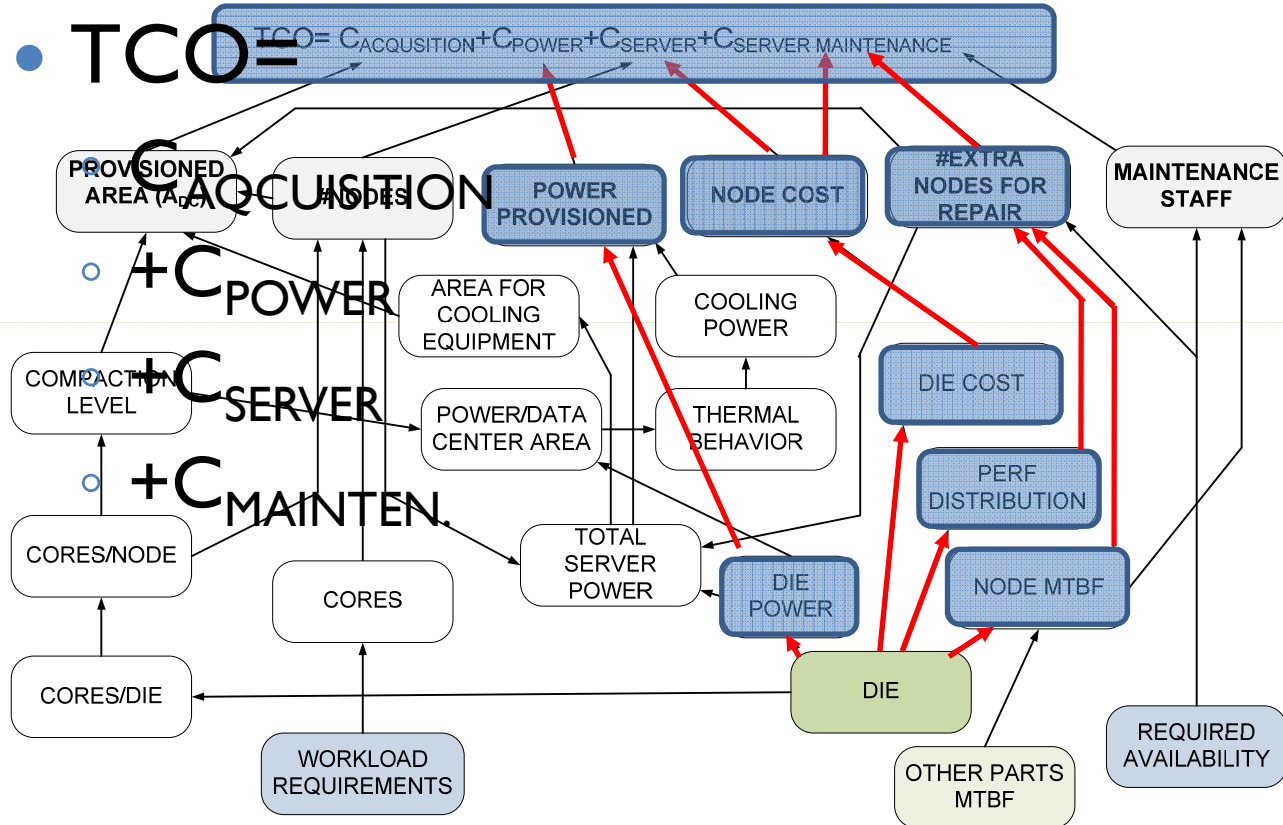


- Scaling -> increased probabilities of failures
- Process Variation, aging, operation below  $V_{ccmin}$  (faults in the field)
- Disabling -> **PERFORMANCE LOSS -> AVAILABILITY LOSS**
- RMT techniques take thread contexts -> Availability loss
- If all (n) nodes @ 80% perf (due to faults), we need 25% more nodes for our initial requirements ( $n \times 100\%$ ).
- all chips @80% perf  $\Leftrightarrow$  80% chips @100% perf
- So, what is the Performance Probability Distribution

Availability = dot product



# Implications of Reliability Decisions on TCO of a Datacenter



Reliability choices

# Implications of Reliability Decisions

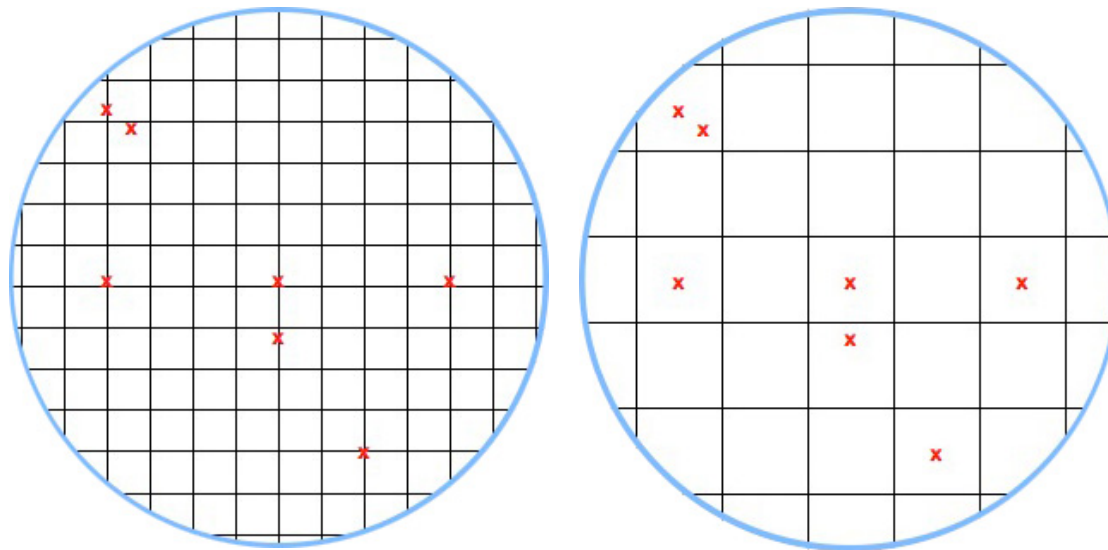
Array/ cell size	1x	1,5x	2x
RegFile			x
Pred	x		
L1Tg		x	
L1D	x		
L2Tg	x		
L2D	x		

- Similar for voltage
- Implications on die size(die cost), power, latency (perf)
- Which combinations minimize TCO?



# Implications of Reliability Decisions

- Increasing SRAM cell size =>
  - more reliable cells
  - but greater **latency** and **power** consumption
  - larger dies => less dies/wafer
    - => **die cost**
    - => lower yield

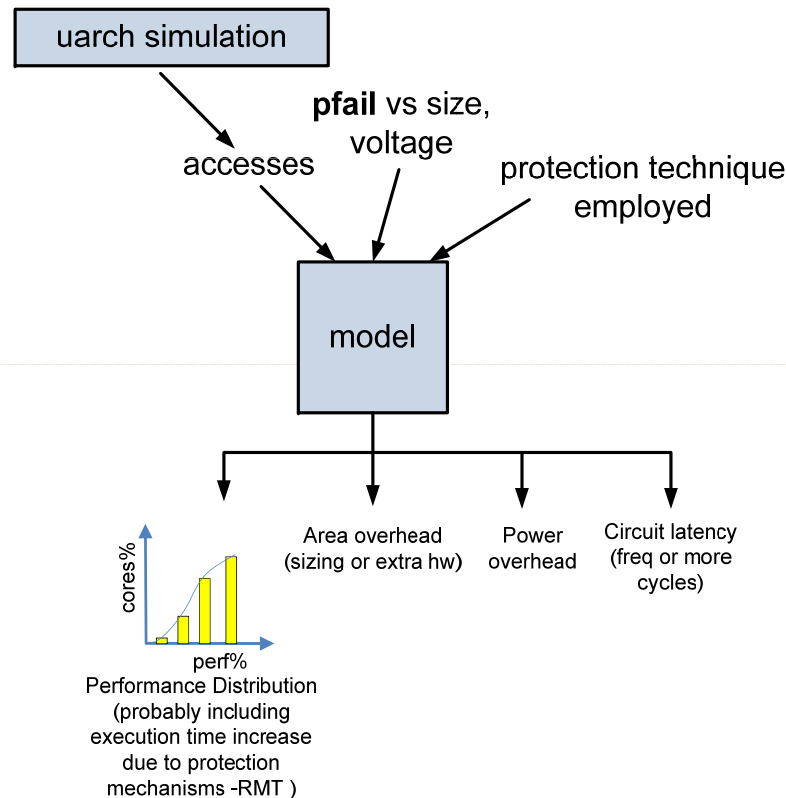


# Implications of Reliability Decisions

- Decreasing voltage => less reliable, but less power consumption
- Some units are critical – single points of failure: need more protection
- Redundant multithreading can take 50% availability.
- Employing complex reliability techniques may increase
  - Design effort
  - Die size (increasing die cost, probability of failure, decreasing yield)



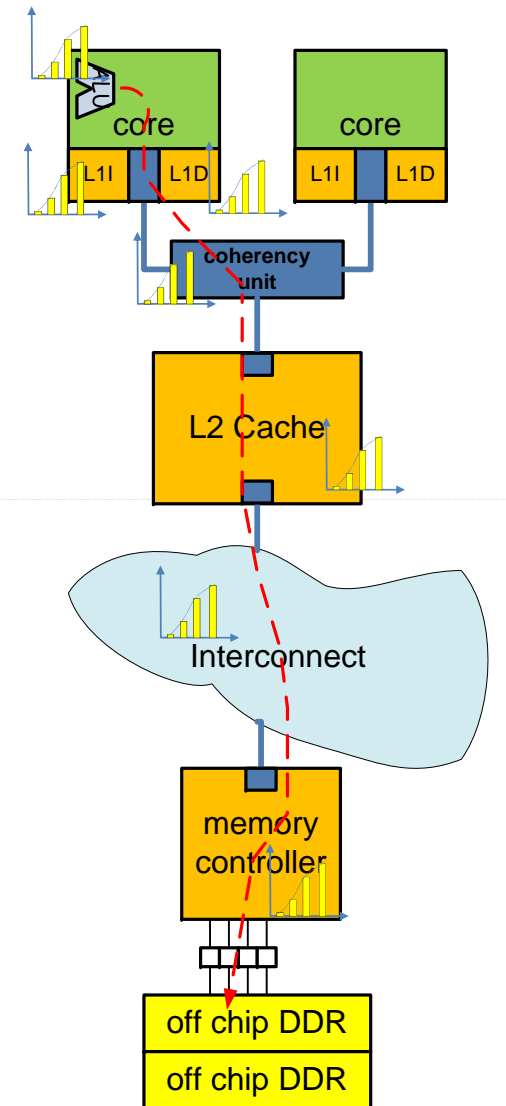
# Analytical Framework - Primitives



- Obtain a Performance Probability distribution for every structure (“primitive”):
  - using accesses (obtained from uarch simulation)
  - without long simulations, fault maps generation

# Analytical Framework

- Characterize all the primitives & combine them.
- Obtain the Global (per chip) Performance Probability Distribution + Power Estimation.
- Estimate TCO
- Explore the Reliability Design Space in less time..



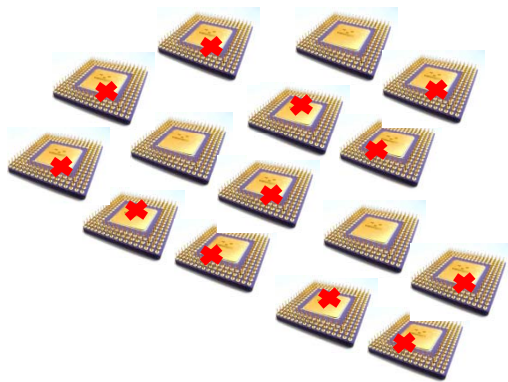
**Thank you!**

---



# Want to find the Perf Distribution of chips— The conventional approach

Let's use 1000 fault maps

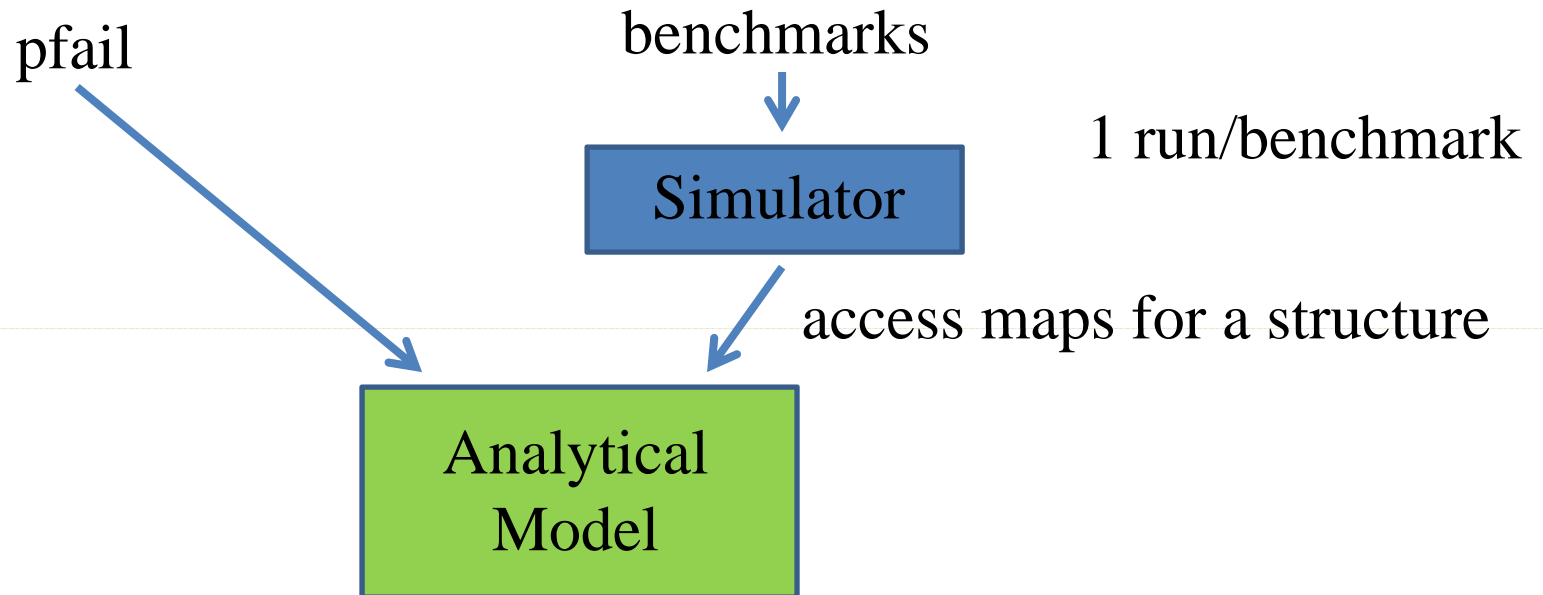


Need  $26 \times 1000 \times 20 \text{min} = 1$  year for simulations (or  $\sim 1.2$  months with 100 PCs)

But, are 1000 fault maps representative?

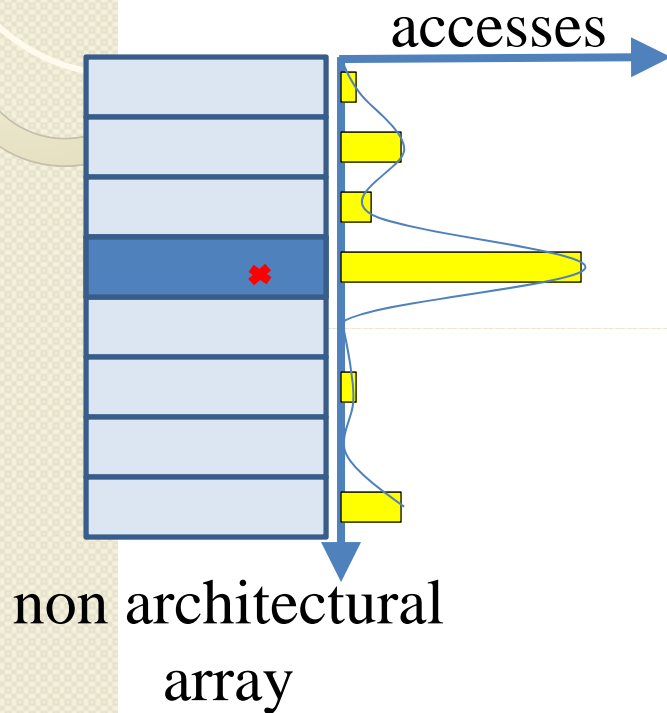
How to explore different configurations?

# Our Approach



- **Accurate Performance Probability Distribution**
  - **analytically**
  - **without long microarchitectural simulations**
  - **without monte carlo simulations**

# Analytical Model



- Given a bit probability of failure **pfail**:
  - Which is the probability that a fault occurs at a hotly accessed position?
  - How much performance is affected in that case?
  - How many chips are affected and how much?
  - => Performance Probability Distribution (binning)
  - not just Yield Analysis

