IEEE/IFIP DSN 2020 Conference – June 29, 2020

Tutorial #1

Cross-Layer Soft-Error Resilience Analysis of Computing Systems

Alberto Bosio
École Centrale de Lyon, France

Stefano Di Carlo and **Alessandro Savino**
Politecnico di Torino, Italy

Dimitris Gizopoulos
University of Athens, Greece

Ramon Canal
Universitat Politècnica de Catalunya
Barcelona Supercomputing Center, Spain

POLITECNICO DI TORINO

TESTGROUP
TEST & RELIABILITY FOR NEXT TECHNOLOGIES

Part #5 Stochastic based approach for System level resilience assessments

Alessandro Savino
Politecnico di Torino, Italy

IEEE/IFIP DSN 2020 Conference – June 29, 2020

OUTLINE

1. Stochastic Models
2. Bayesian Cross-Layer Analysis
3. Bayesian Cross-Layer DSE
4. Conclusions and perspectives

OUTLINE

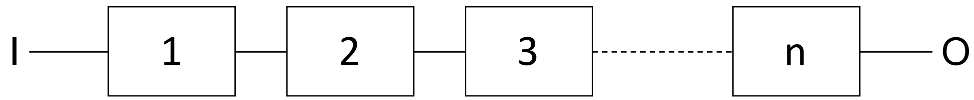
1. Stochastic Models
2. Bayesian Cross-Layer Analysis
3. Bayesian Cross-Layer DSE
4. Conclusions and perspectives

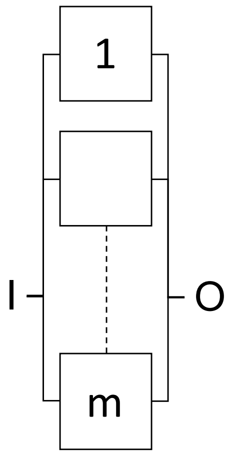
Reliability Block Diagrams

Basic Concepts

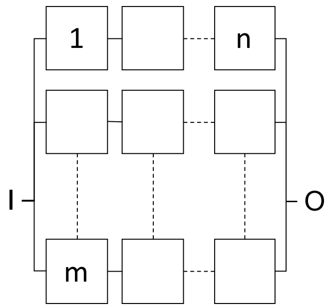
- RBD is a graphical representation of the arrangement of system components regarding their influence on the system reliability.
- An RBD demonstrates the effect of the success or failure of a component on the success or failure of the whole system.

Series: all components must be operational

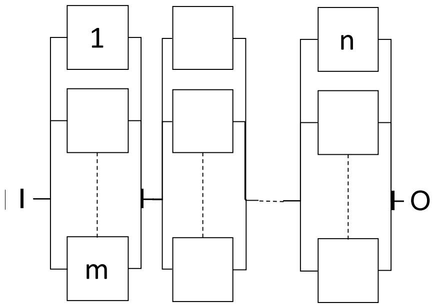




Parallel: at least one of the components must be operational



parallel-series



series-parallel

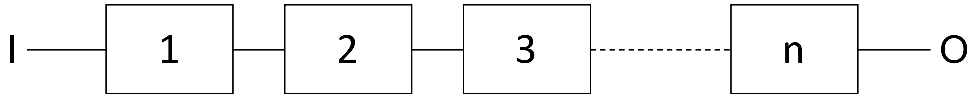
Reliability Block Diagrams

Evaluate the Resiliency

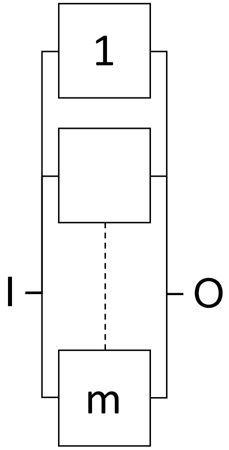
- Given a survival function $R_i(t)$ for each block:
  - $Pr(X_i(t) = \text{Fault Free}) = R_i(t)$
  - $Pr(X_i(t) = \text{Faulty}) = 1 - R_i(t)$
- Example:
  - How to evaluate Mean Time To Failure:

$$MTTF = \int_{t=0}^{\infty} R_s(t)dt$$

$$R_s(t) = \prod_{i=1}^{n} R_i(t)$$



$$R_s(t) = 1 - \prod_{i=1}^{m} [1 - R_i(t)]$$

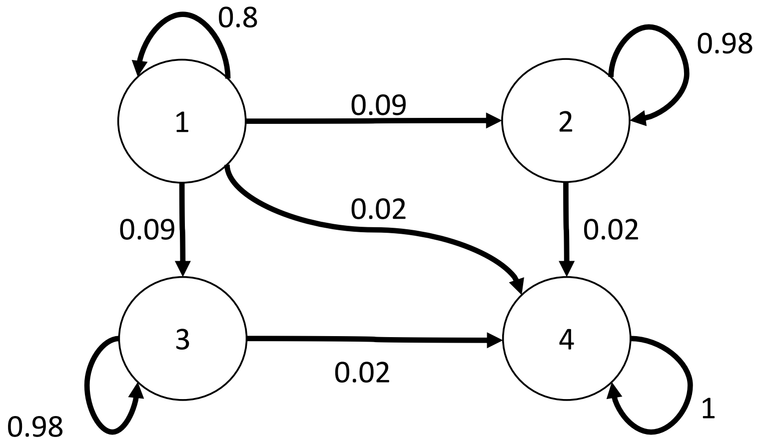


Markov Chains

Basic System Composition

- Markov chains (MCs) are stochastic processes whose futures are conditionally independent of their pasts given their present values.
- In practical terms, for each component of the system is accounted with all its possible states to compose a Transition Diagram describing the Full System Reliability behavior.

State	Component 1	Component 2
1	Fault-Free	Fault-Free
2	Fault-Free	Faulty
3	Faulty	Fault-Free
4	Faulty	Faulty

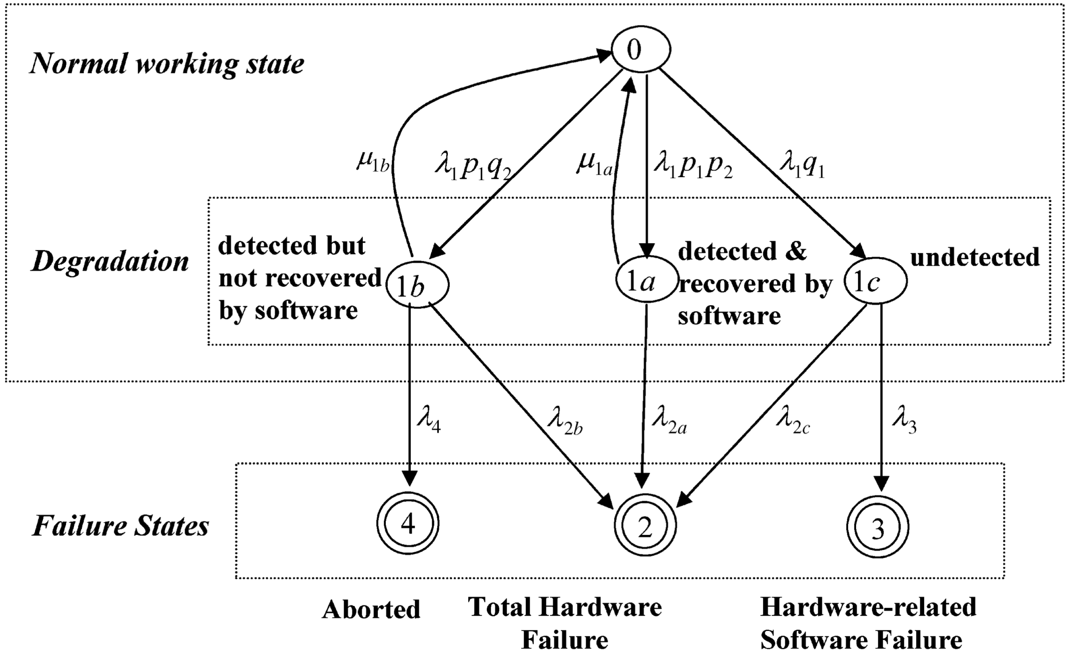


$$P_{x_i}(t+1) = \sum_{k=0}^{n-1} P_{k,i}(t) \cdot P_{x_k}(t)$$

Markov Chains

Cross-Layer Resilience Evaluation

- A Cross-Layer Model must:
  1. Enumerate all possible HW faults
  2. Enumerate all SW states that might happen due to an HW fault
  3. Enumerate all Failure/NonFailure states of the system
- Then, characterize each state with proper transitions (and probability of moving along the transition) in order to define the full model.



Picture from: Teng X, Pham H, Jeske DR. Reliability Modeling of Hardware and Software Interactions, and Its Applications. IEEE Transactions on Reliability. 2006 Dec;55(4):571–577

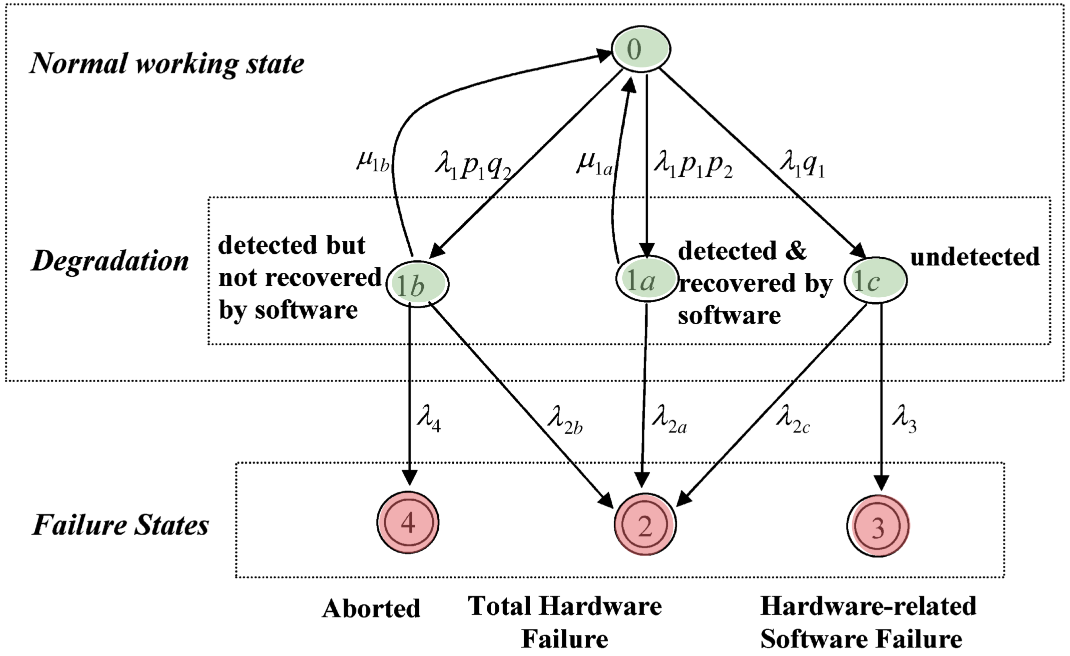
Markov Chains

Cross-Layer Resilience Evaluation

- The Resilience can then be evaluated by calculation of the Probability of being in all "fault-free" states at time t.

$$R(t) = P_0(t) + P_{1a}(t) + P_{1b}(t) + P_{1c}(t)$$

$$MTBF = \frac{1}{1 - R(t)}$$

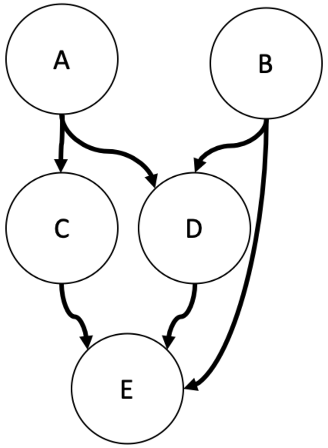


Picture from: Teng X, Pham H, Jeske DR. Reliability Modeling of Hardware and Software Interactions, and Its Applications. IEEE Transactions on Reliability. 2006 Dec;55(4):571–577

Bayesian Networks

Basic Concepts

- Bayesian Networks (BN) are stochastic models, defined as a Directed Acyclic Graph (DAGs) where:
  - Nodes are random variables (say components) that can assume different states.
  - Edges are the relationship of dependence between nodes (destination depends from its sources)



θ_D	A-1		A-2		A-3	
	B-1	B-2	B-1	B-2	B-1	B-2
D-1	0.5	0.2	0	0	0.3	0
D-2	0	0.6	0.7	0	0.3	0.5
D-3	0.25	0.1	0	1	0.3	0.5
D-4	0.25	0.1	0.3	0	0.1	0

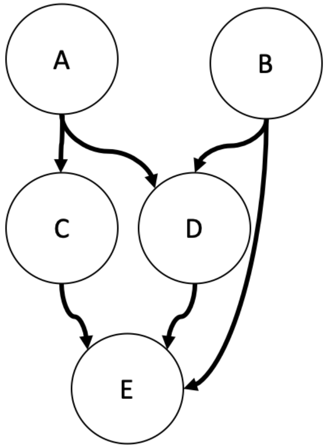
$P(D=D\text{-}2 \mid A=A\text{-}2, B=B\text{-}1)$

Each node is provided by a Conditional Probability Table (CPT): it defines the conditional probabilities of the node with respect to the state of its parents. When the node has no parents the CPT represents the marginal probability distribution for the state.

Bayesian Networks

Evaluation

- The evaluation of a Bayesian network is the computation of the probabilities for each node of being in every possible states.



θ_D	A-1		A-2		A-3	
	B-1	B-2	B-1	B-2	B-1	B-2
D-1	0.5	0.2	0	0	0.3	0
D-2	0	0.6	0.7	0	0.3	0.5
D-3	0.25	0.1	0	1	0.3	0.5
D-4	0.25	0.1	0.3	0	0.1	0

P(D=D-2 | A=A-2,B=B-1)

$$P(n_1, n_2, \cdots, n_m) = \prod_{i=1}^{m} P(n_i | parents(n_i))$$

Bayesian Network Belief Update

Pros and Cons

What is best for Cross-Layer Resilience Evaluation?

- RBDs
  - Pros: Component based modeling
  - Cons: Too simple description of the propagation
- MCs
  - Pros: Recovery policies as part of the modeling
  - Cons: System State based modeling, limited time dependence
- BNs
  - Pros: Component based modeling, complex chaining of events
  - Cons: no easy recovery mechanisms management due to DAG properties

BAYESIAN CROSS-LAYER Tools

A full framework from components characterization to system optimization

Component Characterization

- Technology Characterization Tools
- CPU/GPU Characterization Tools
- Software Characterization Tools

System Reliability Analyzer (SyRA)

Scan me

[Vallero et al. TOC'18]

Reliability Design Optimizer (ReDO)

Scan me

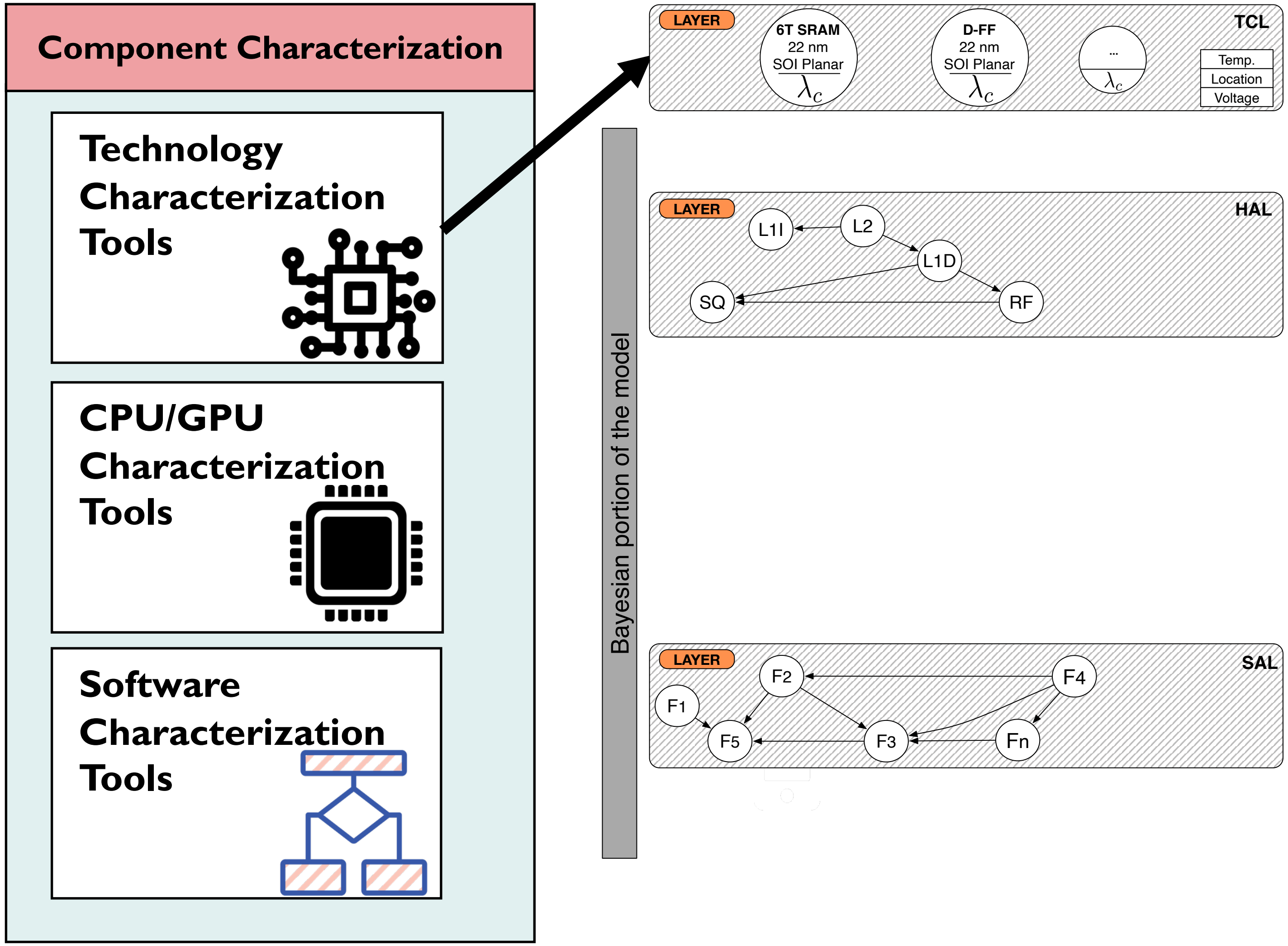
[Savino et al. TOC'18]

OUTLINE

1 Stochastic Models
2 Bayesian Cross-Layer Analysis
3 Bayesian Cross-Layer DSE
4 Conclusions and perspectives

BAYESIAN CROSS-LAYER ANALYSIS

System Reliability Analyzer (SyRA)



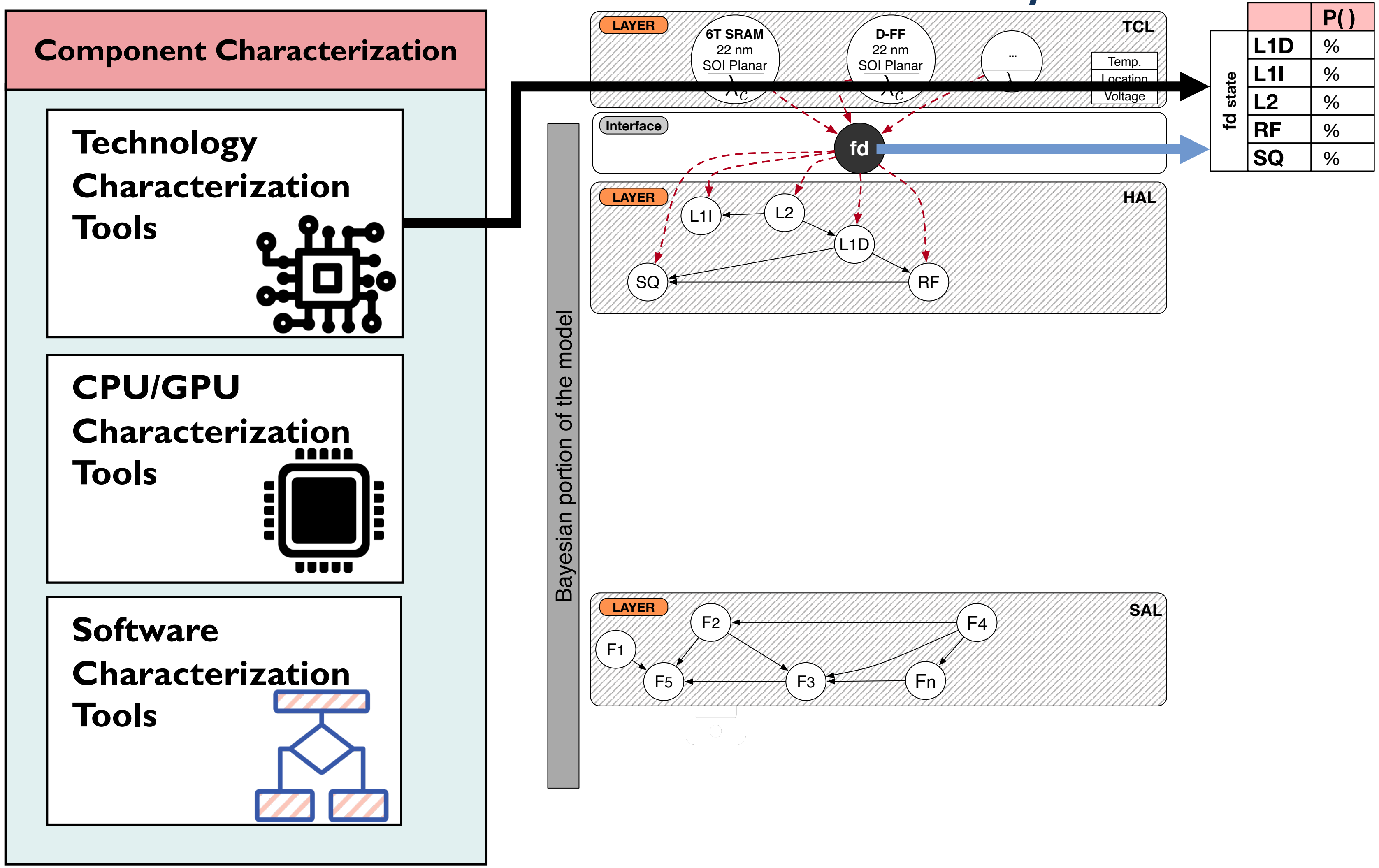
The core of the model is a Bayesian network that represents how soft errors propagate through the layers of the system

The output of the model is a set of resiliency metrics for the system computed by means of Bayesian inference.

[Vallero et al. TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

System Reliability Analyzer (SyRA)



Reliability model organized into layers connected through a set of special macro nodes called interfaces

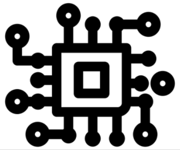
[Vallero et al. TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

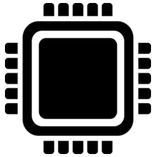
System Reliability Analyzer (SyRA)

Component Characterization

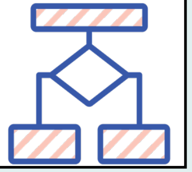
Technology Characterization Tools

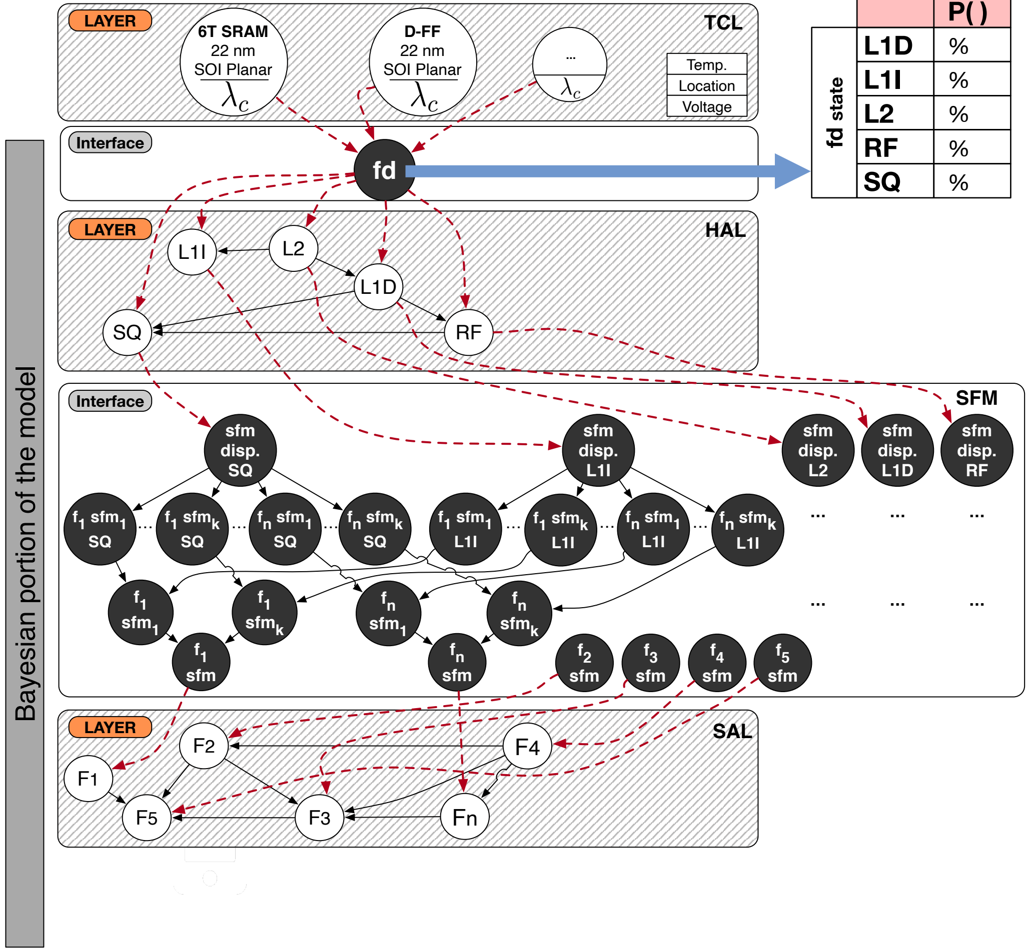


CPU/GPU Characterization Tools



Software Characterization Tools





fd state		P( )
	L1D	%
	L1I	%
	L2	%
	RF	%
	SQ	%

Reliability model organized into layers connected through a set of special macro nodes called interfaces

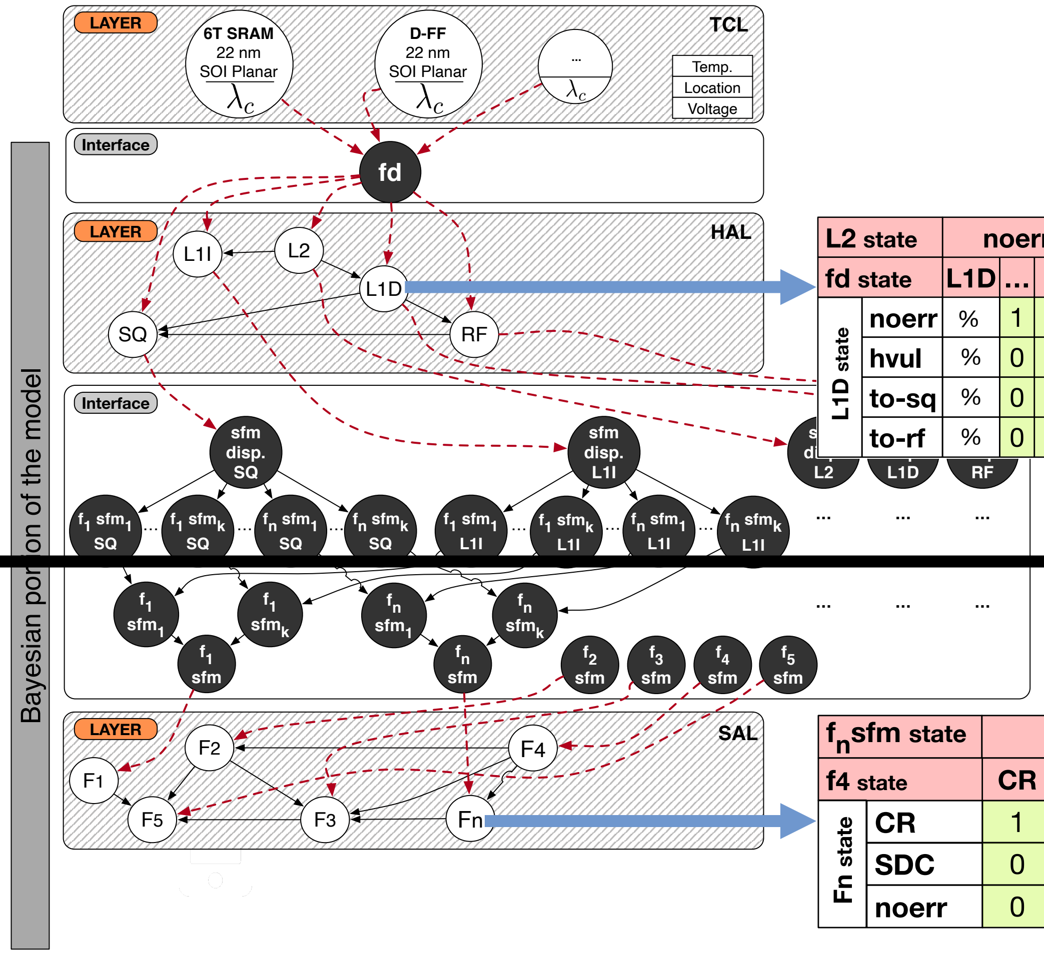
[Vallero et al. TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

System Reliability Analyzer (SyRA)

Component Characterization

- Technology Characterization Tools
- CPU/GPU Characterization Tools
- Software Characterization Tools



L2 state	noerr			hvul			to-l1d			to-l1i		
fd state	L1D	…	L2	L1D	…	L2	L1D	…	L2	L1D	…	L2
L1D state: noerr	%	1	1		…	1		…	%			1
L1D state: hvul	%	0	0		…	0		…	%			0
L1D state: to-sq	%	0	0		…	0		…	%			0
L1D state: to-rf	%	0	0		…	0		…	%			0

$f_n$sfm state	noerr			sfm$_1$			sfm$_2$	…
f4 state	CR	SDC	noerr	CR	SDC	noerr		
Fn state: CR	1	%	0	1	%	%		
Fn state: SDC	0	%	0	0	%	%		
Fn state: noerr	0	%	1	0	%	%		

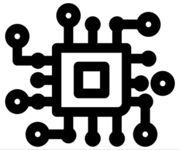
[Vallero et al. TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

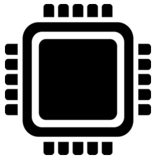
System Reliability Analyzer (SyRA)

Component Characterization

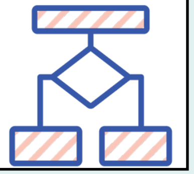
Technology Characterization Tools

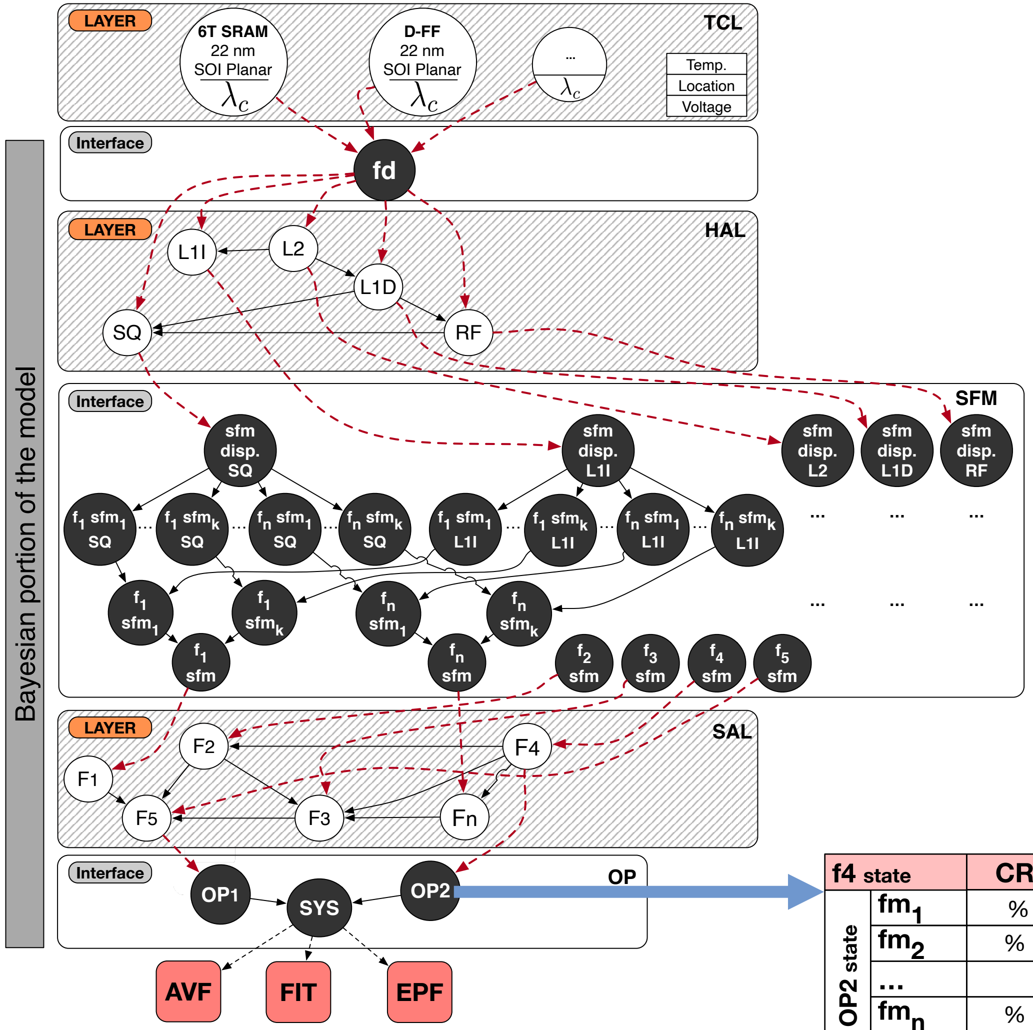


CPU/GPU Characterization Tools



Software Characterization Tools





	f4 state	CR	SDC	mask
OP2 state	fm_1	%	%	%
	fm_2	%	%	%
	…			
	fm_n	%	%	%
	noerr	%	%	%

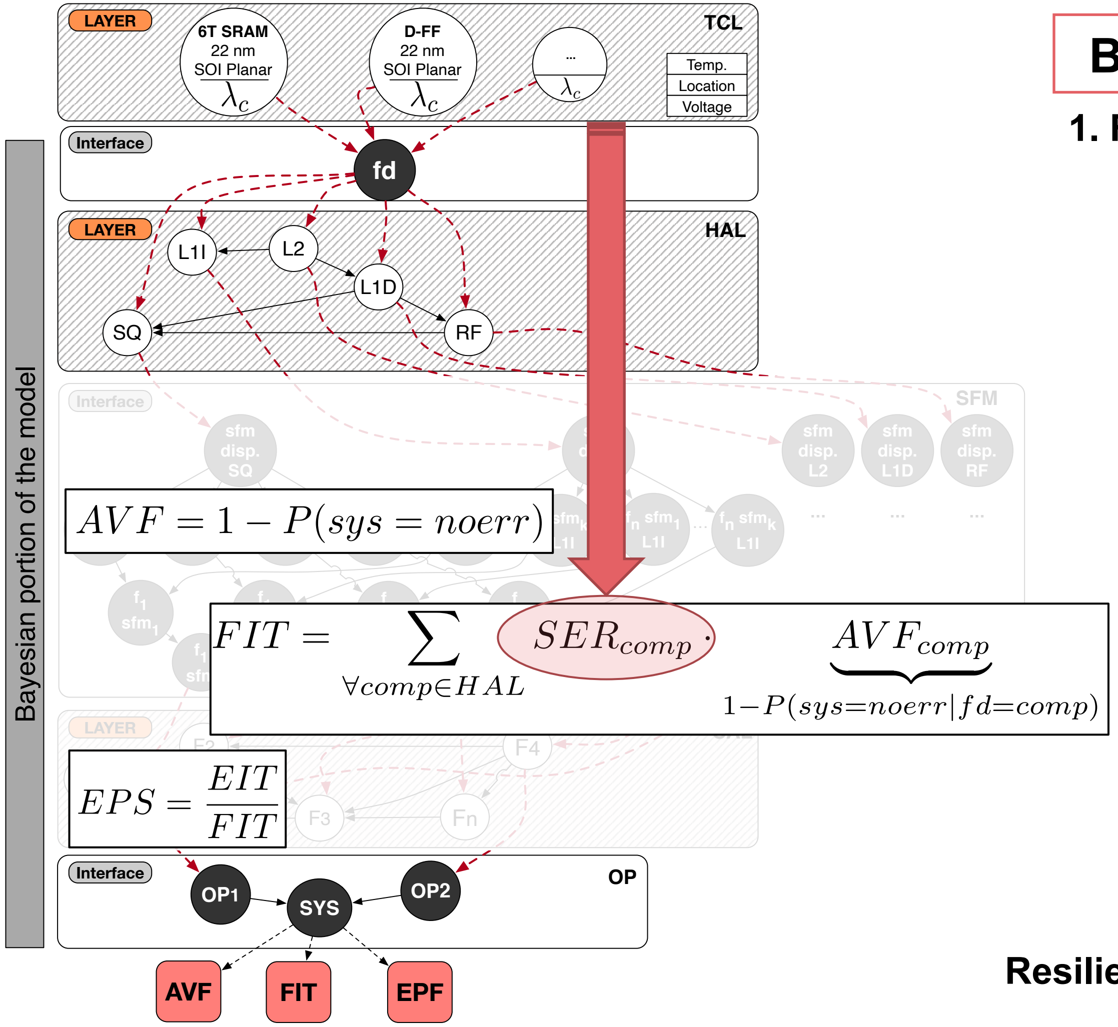
[Vallero et al. TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

System Reliability Analyzer (SyRA)

Component Characterization

- **Technology Characterization Tools**
- **CPU/GPU Characterization Tools**
- **Software Characterization Tools**



$$AVF = 1 - P(sys = noerr)$$

$$FIT = \sum_{\forall comp \in HAL} SER_{comp} \cdot \underbrace{AVF_{comp}}_{1-P(sys=noerr|fd=comp)}$$

$$EPS = \frac{EIT}{FIT}$$

[Vallero et al. TOC'18]

Bayesian Reasoning

1. Failure rate

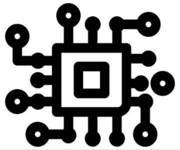
Resiliency metrics

BAYESIAN CROSS-LAYER ANALYSIS

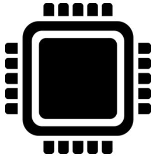
System Reliability Analyzer (SyRA)

Component Characterization

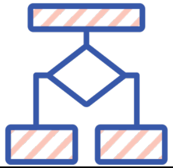
Technology Characterization Tools

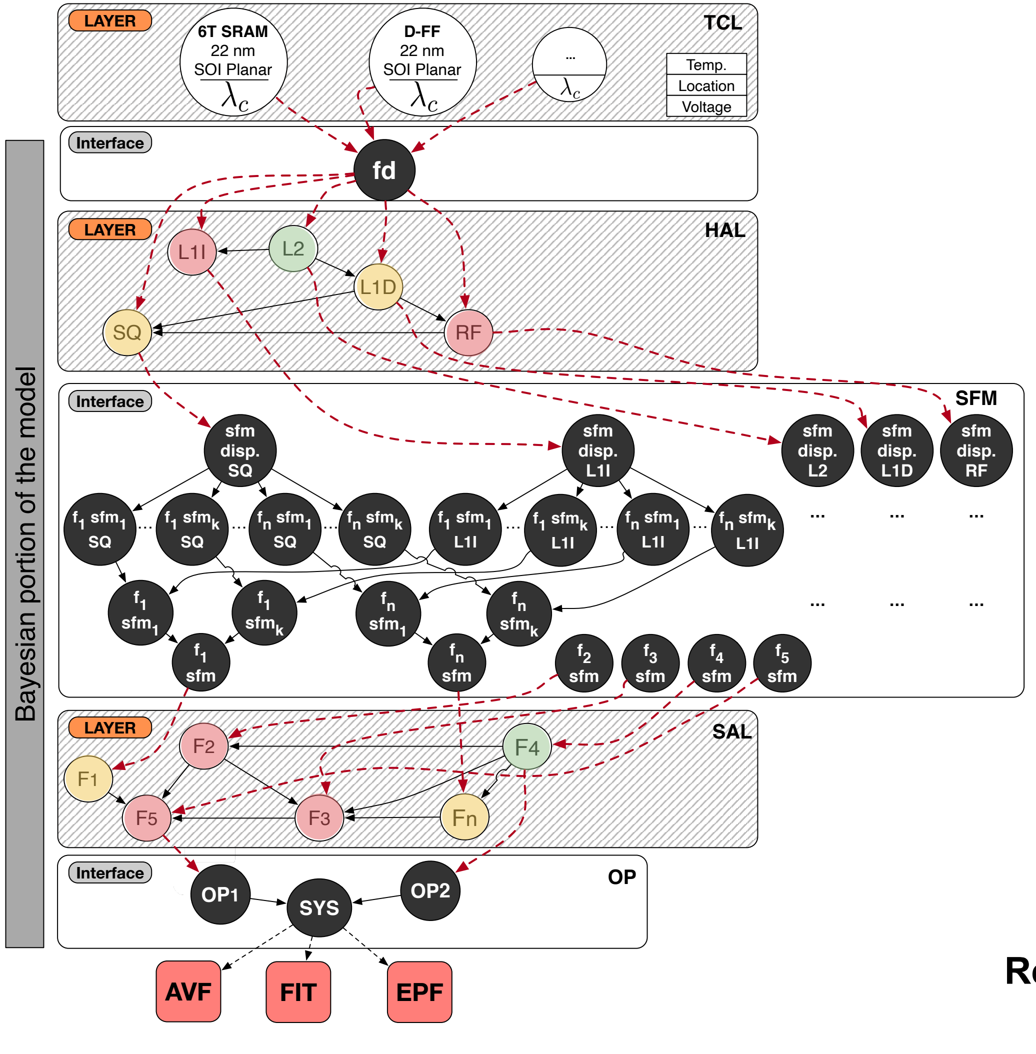


CPU/GPU Characterization Tools



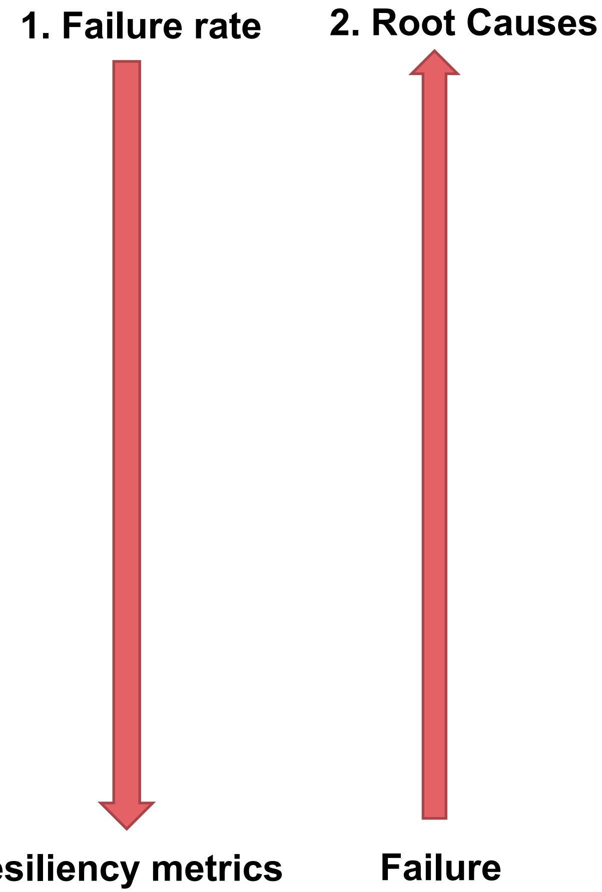
Software Characterization Tools





[Vallero et al. TOC'18]

Bayesian Reasoning



BAYESIAN CROSS-LAYER ANALYSIS

Benchmark Applications & Systems

1. MiBench
   1. ss: Stringsearch
   2. s_(x): Susan (smoothing, edges, corners)
   3. aes: AES
   4. qs: Qsort
   5. ff: FFT
   6. sha: Sha
2. Custom
   1. rta: control application from the avionic domain (IP)
   2. bm: bare metal application for DC motor controllers (IP)
   3. hpc: open source software to solve hyperbolic equations on dynamically changing fully-adaptive conforming 2D triangular grids (from: https://www5.in.tum.de/sierpinski/index.php)

ARM Cortex-A9

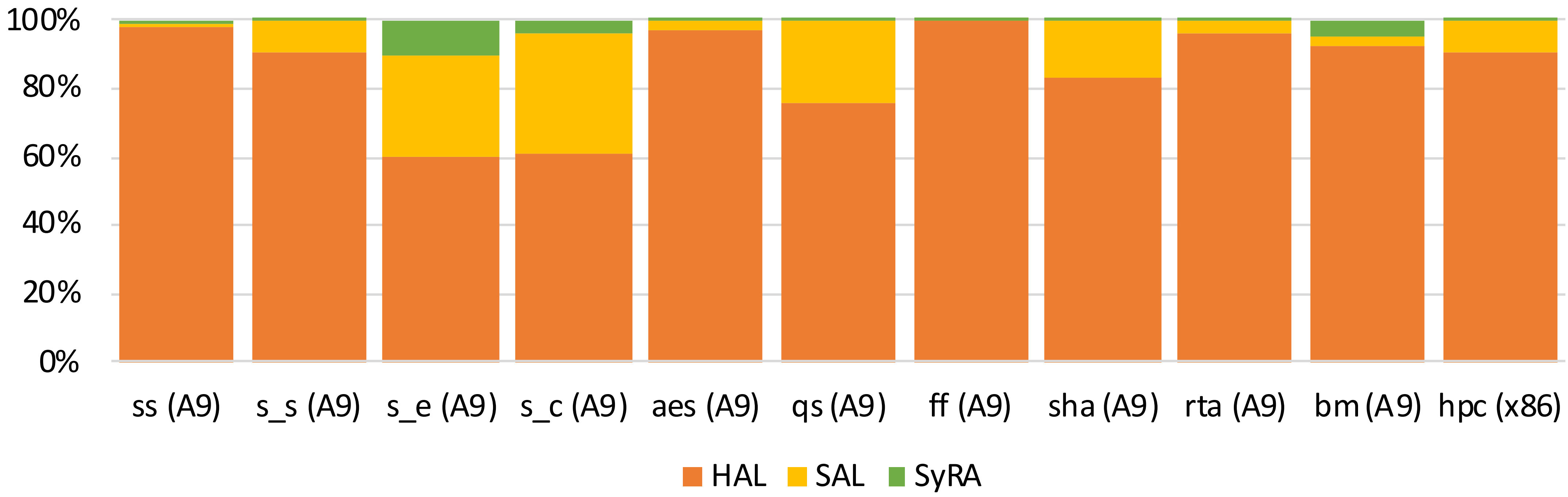
intel SKYLAKE

	A9	A15	X86
L1 I/D	32KB	32KB	32KB
L2	512KB	1MB	1MB
RF	56 32bit reg	128 32bit reg	168 64bit reg
SQ	8 32bit reg	16 32bit reg	72 64bit reg.
Tech. Node	65/45nm	32/28nm	14 nm FinFET
Clock	0.8-2GHz	1-2GHZ	4GHz

[Vallero et al. TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

BN Building and Analysis Time

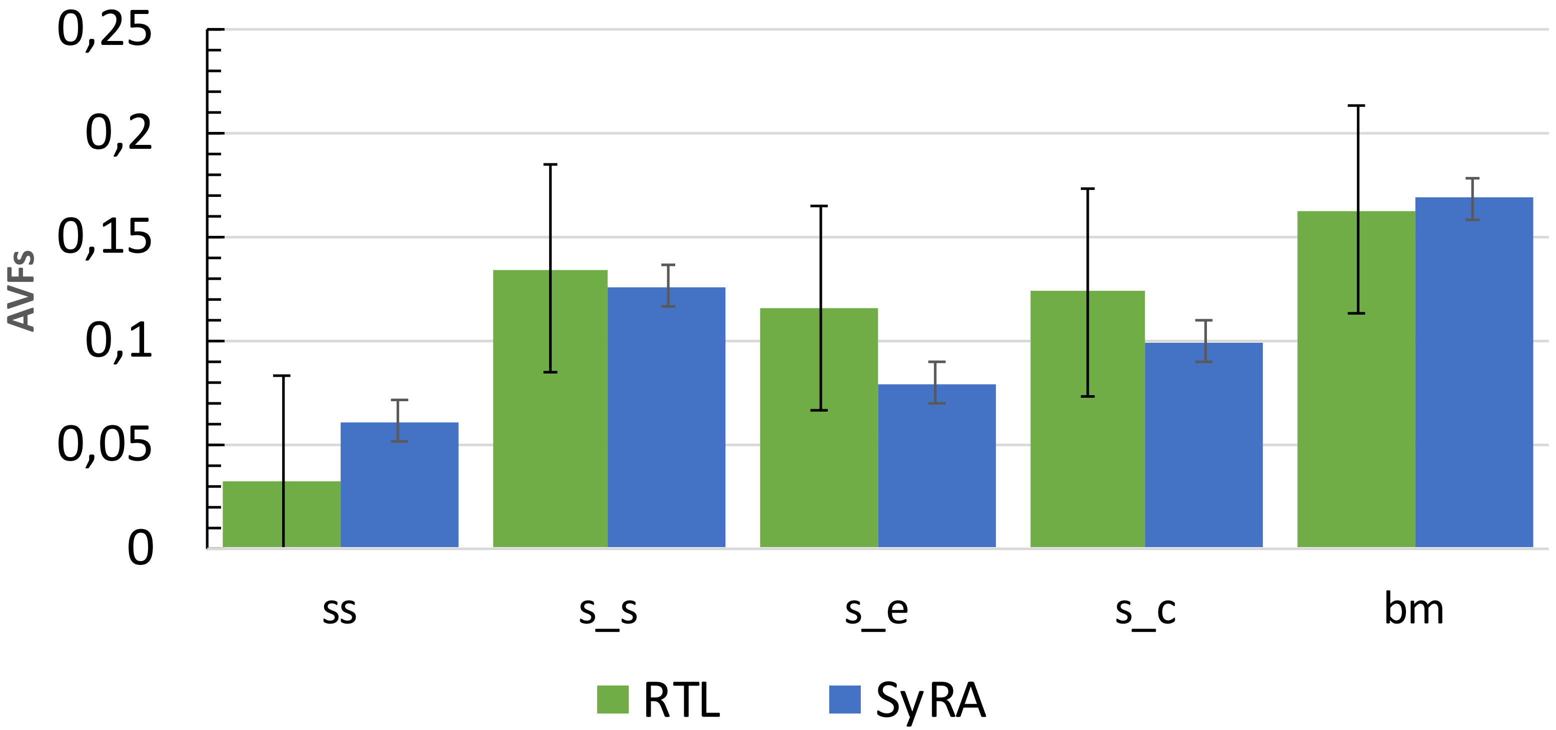


[Vallero et al. TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

RTL fault injection vs BN (Accuracy)

ARM Cortex-A9



~4000 RTL injections per block

2% error margin

99% CI

[Chatzidimitriou et al. DSN'17] [Vallero et. al IEEE TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

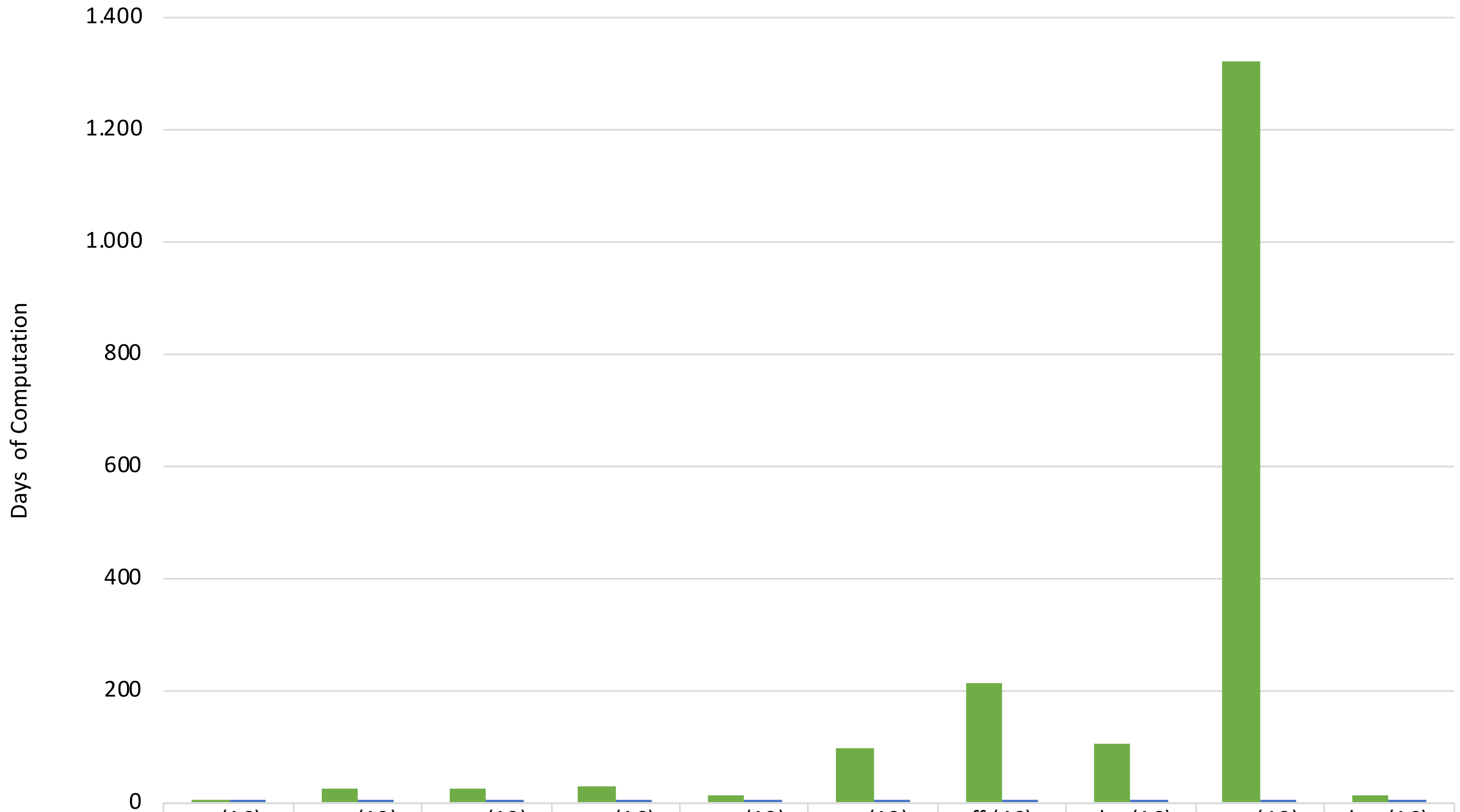
RTL fault injection vs BN Time

ARM
Cortex-A9

~4000 RTL injections per block

2% error margin

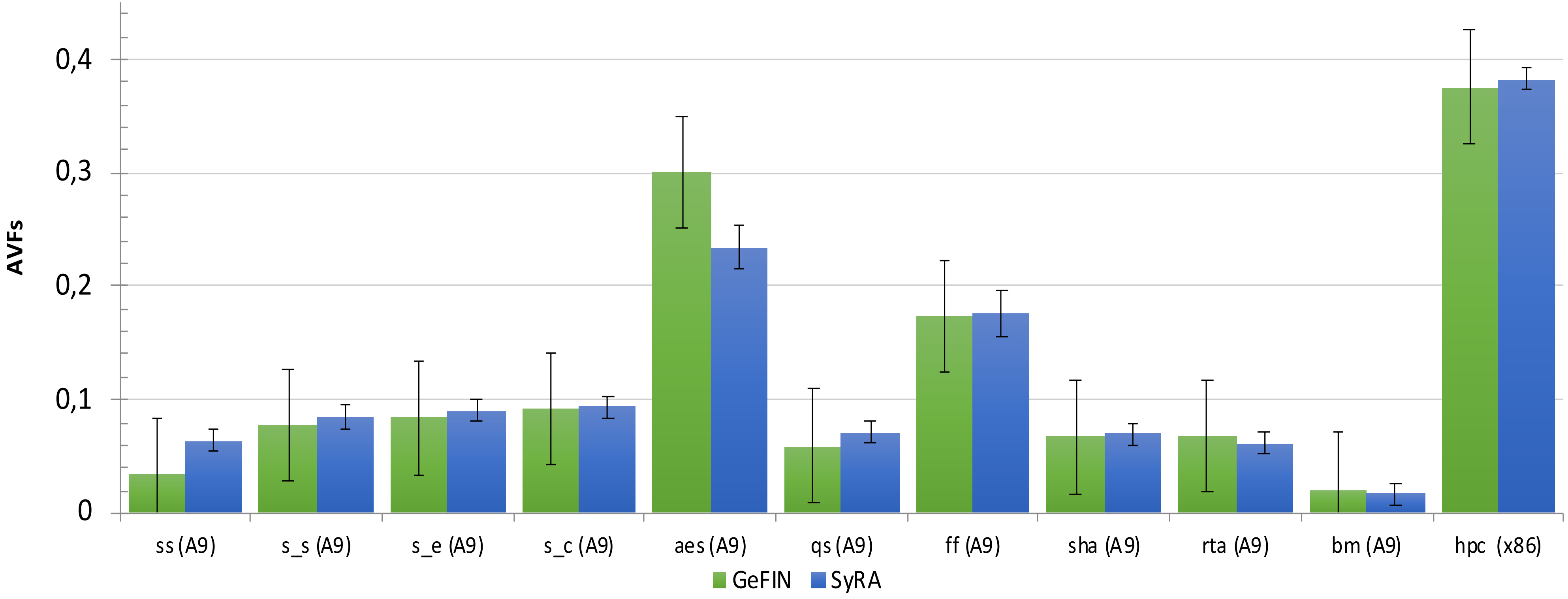
99% CI



	ss (A9)	s_s (A9)	s_e (A9)	s_c (A9)	aes (A9)	qs (A9)	ff (A9)	sha (A9)	rta (A9)	bm (A9)
RTL (days)	1,83	27,32	26,73	31,16	12,63	96,54	214,08	104,61	1320,98	13,42
SyRA (days)	0,08	0,59	0,17	0,14	1,00	0,37	1,47	0,28	1,17	0,09

BAYESIAN CROSS-LAYER ANALYSIS

Microarchitectural Level Fault Injection vs BN (Accuracy)



~2000 uA injections per block

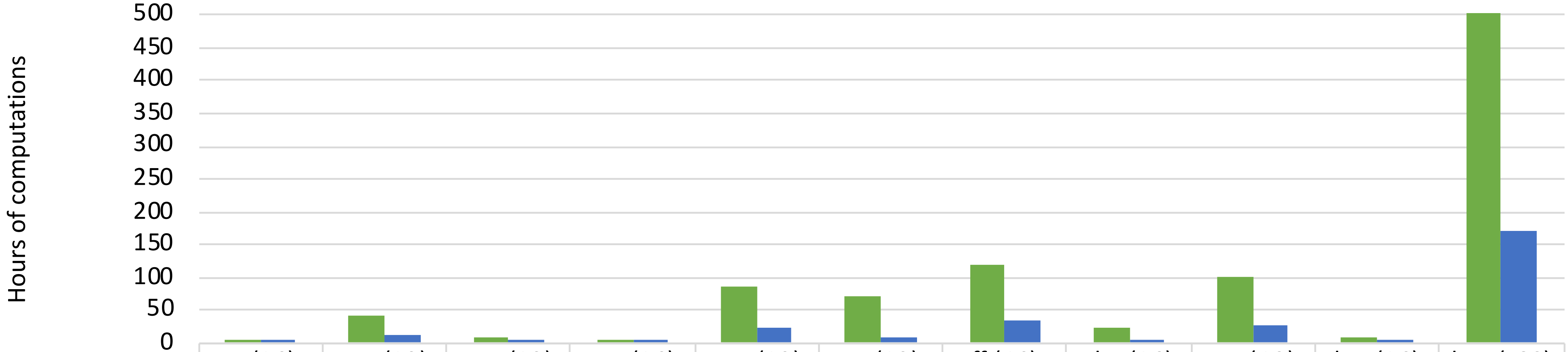
5% error margin

99% CI

[Vallero et al. ITC'16] [Vallero et. al IEEE TOC'18]

BAYESIAN CROSS-LAYER ANALYSIS

Microarchitectural Level Fault Injection vs BN Time



	ss (A9)	s_s (A9)	s_e (A9)	s_c (A9)	aes (A9)	qs (A9)	ff (A9)	sha (A9)	rta (A9)	bm (A9)	hpc (x86)
GeFIN (hours)	5	42	9	7	87	70	119	25	101	8	585
SyRA (hours)	2	14	4	3	24	9	35	7	28	2	172

~2000 uA injections per block

5% error margin

99% CI

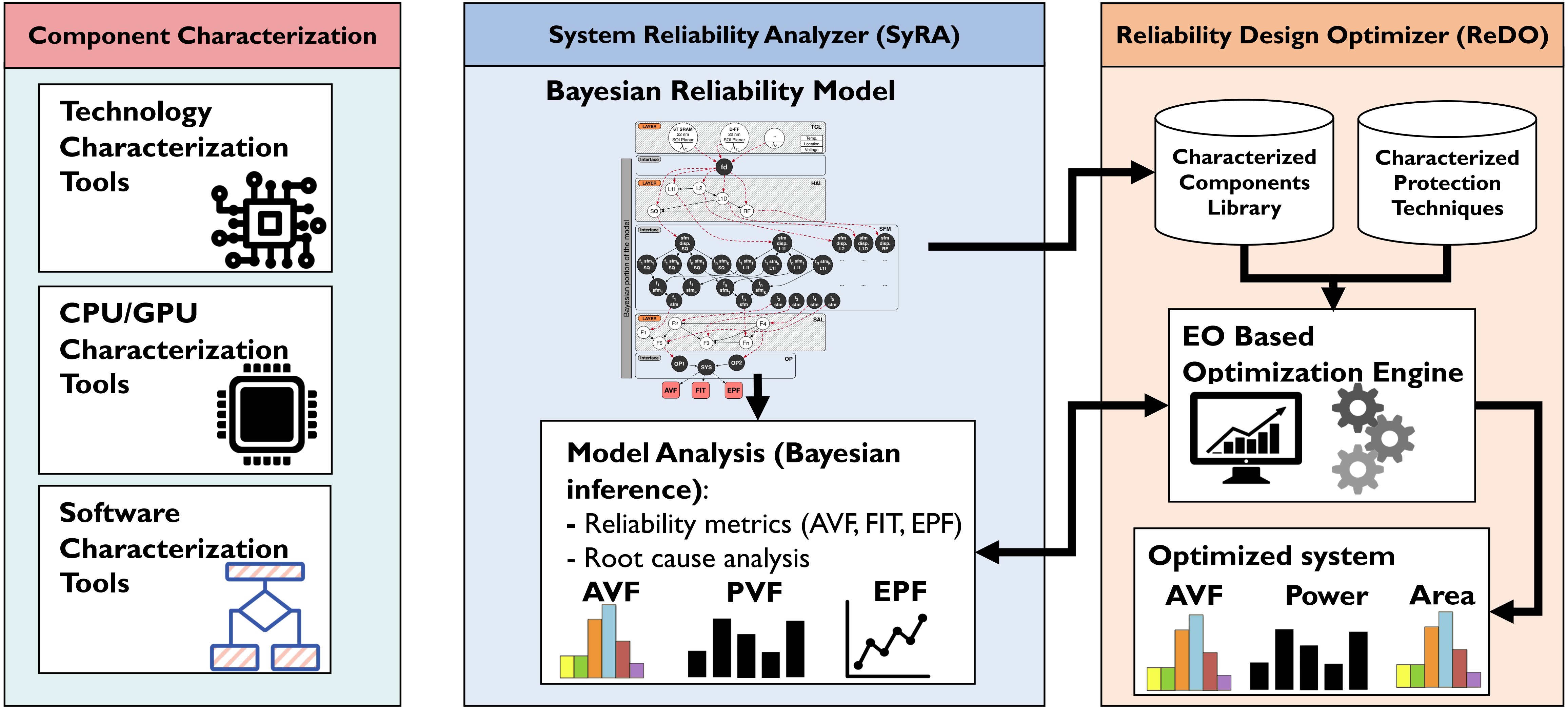
[Vallero et al. TOC'18]

OUTLINE

1. Stochastic Models
2. Bayesian Cross-Layer Analysis
3. Bayesian Cross-Layer DSE
4. Conclusions and perspectives

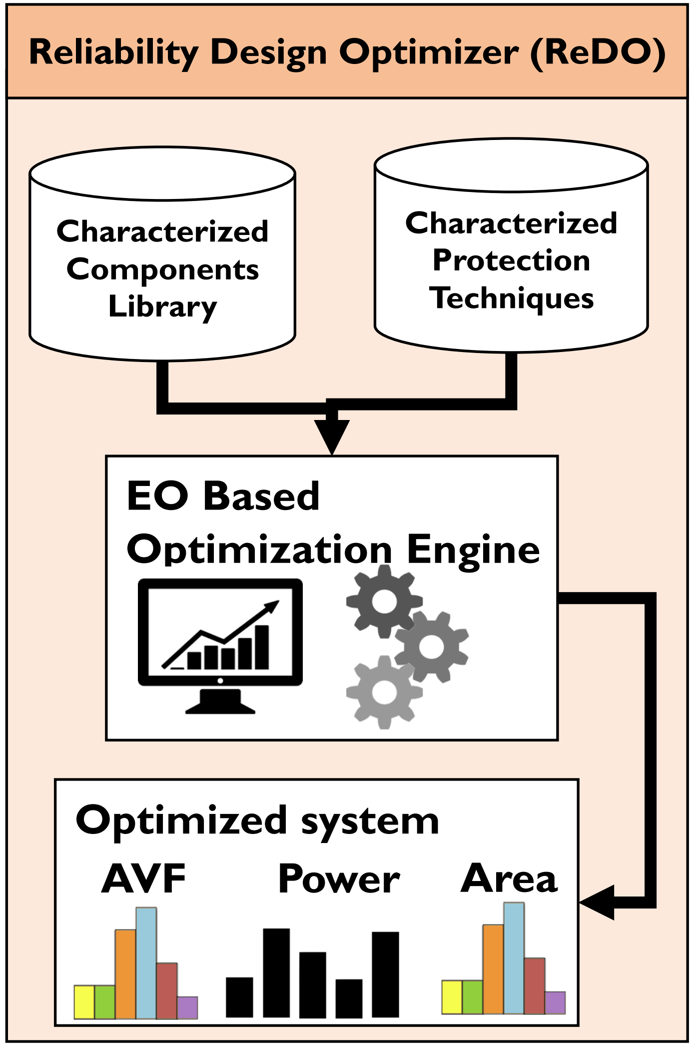
BAYESIAN CROSS-LAYER Tools

A full framework from components characterization to system optimization



BAYESIAN CROSS-LAYER DSE

Reliability Design Optimizer (ReDO)



Extremal Optimization (EO) is an evolutionary optimization heuristic that evolves a single solution through local modifications (no need to store large populations of possible solutions)

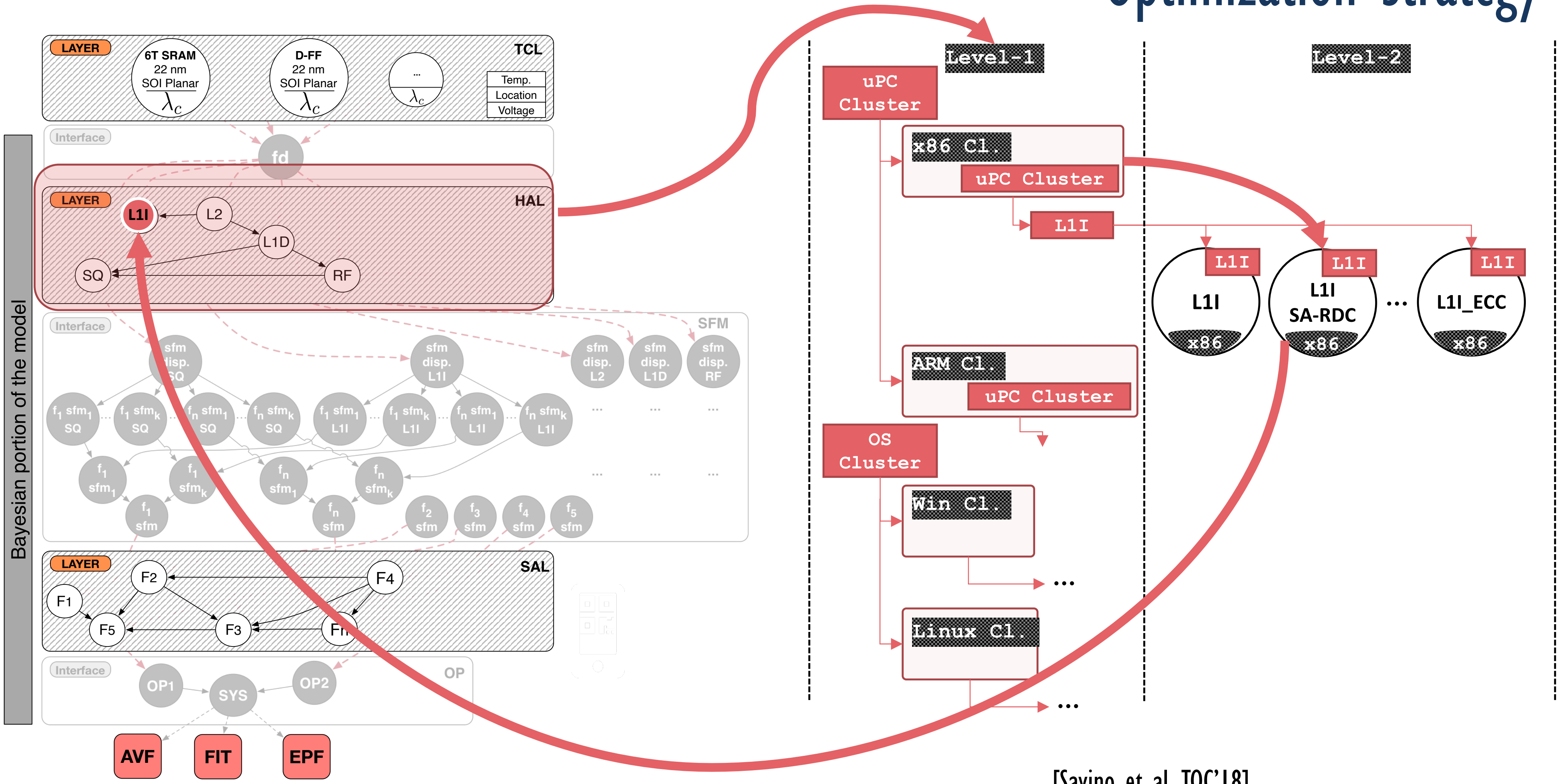
Local modifications remove low-quality components, replacing them with randomly selected new components

Takes advantage of the component-based Bayesian Reliability Model to identify weak components without long simulations (Bayesian Inference)

Combines reliability metrics with other design parameters (e.g., power consumption and area) to enable multi-objective optimization

BAYESIAN CROSS-LAYER DSE

Optimization Strategy



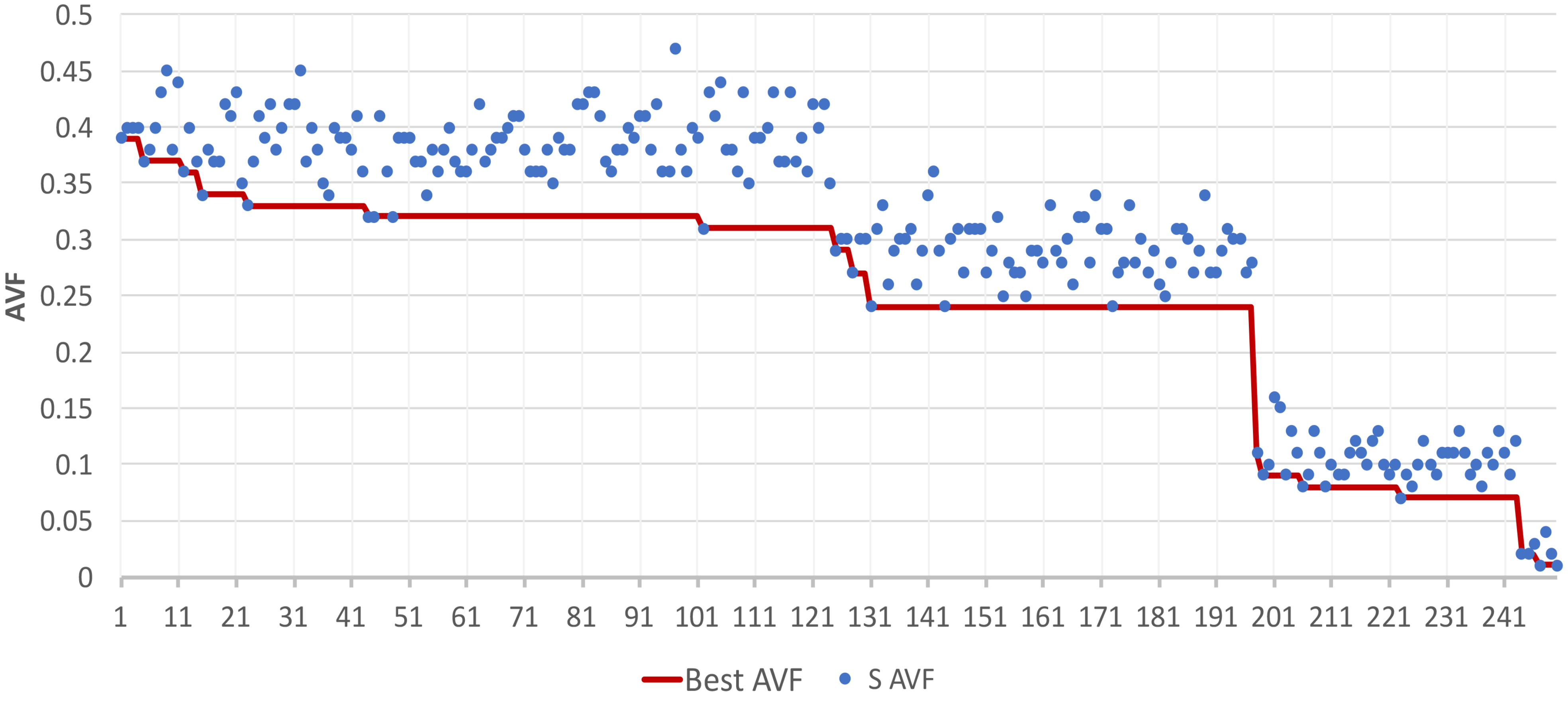
[Savino et al. TOC'18]

BAYESIAN CROSS-LAYER DSE

Trajectory of design space exploration

HPC app

Linux



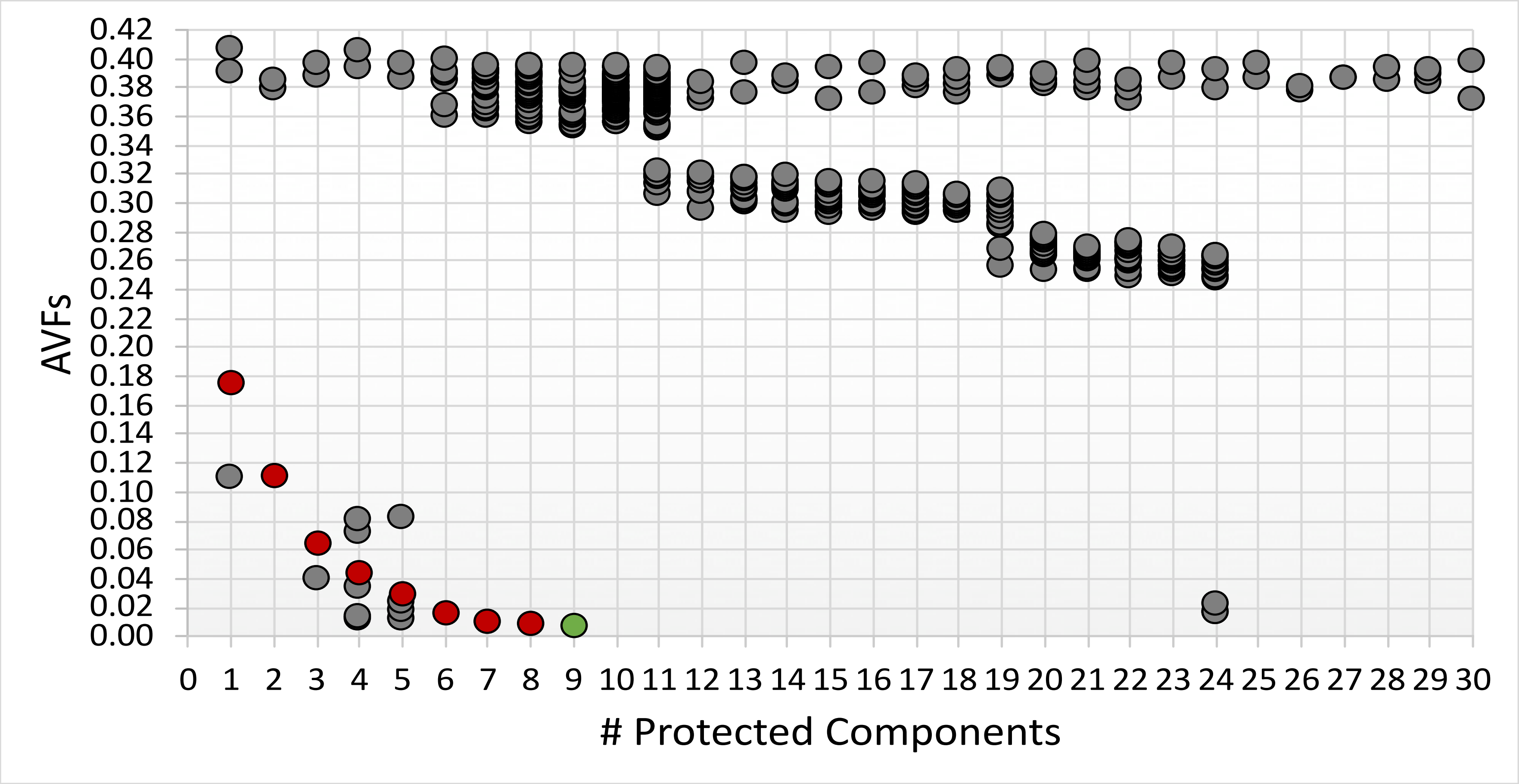
[Savino et al. IEEE TOC'18] AVF trajectory when performing DSE for best AVF.

BAYESIAN CROSS-LAYER DSE

Multi objective-design space exploration

HPC app

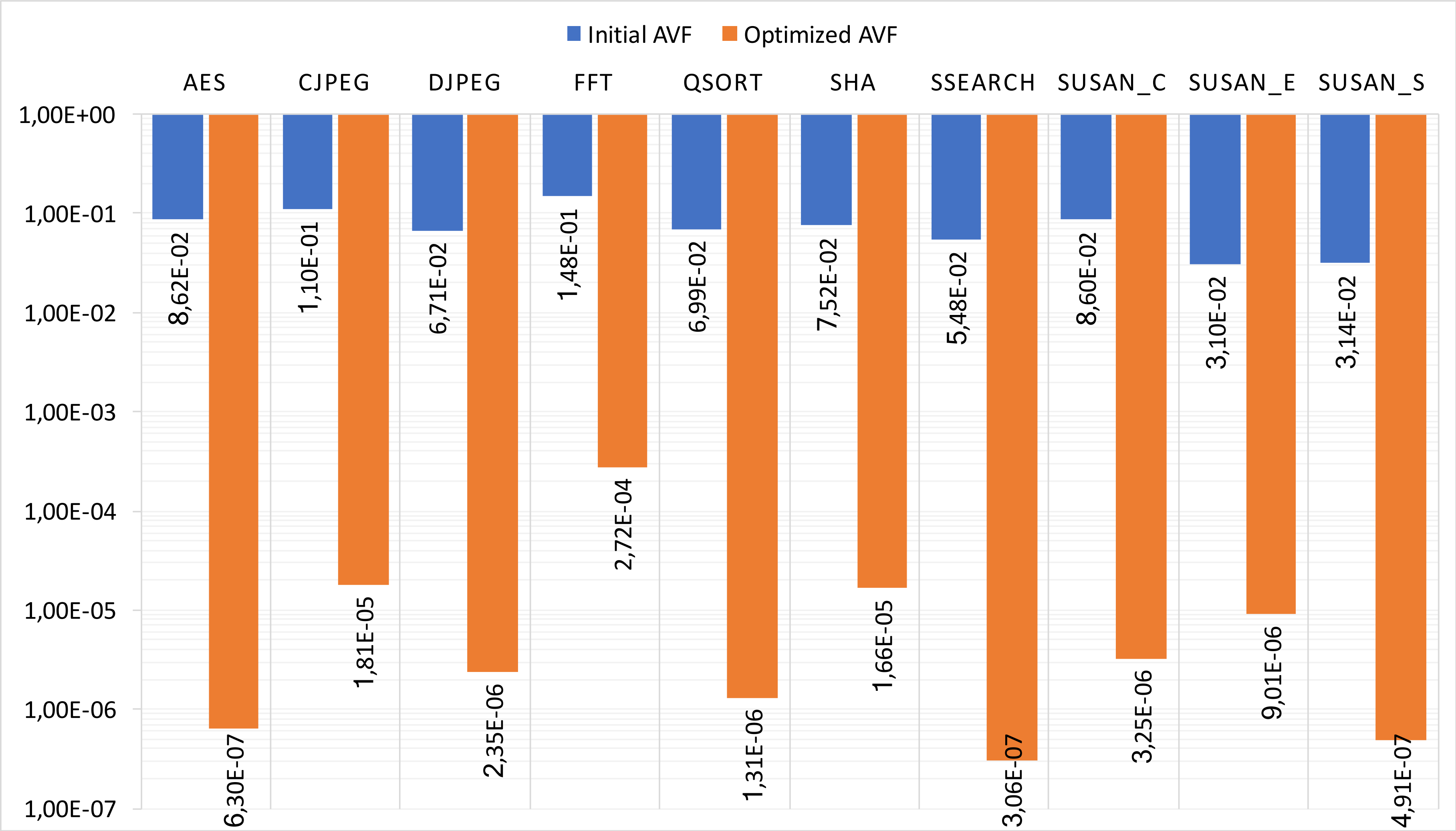
Linux



[Vallero et al. IEEE TOC'18] [Savino et al. IEEE TOC'18] SyRA diagnostic analysis for the hpc benchmark. Gray dots show the AVFs of 600 combinations of randomly protected components. Red dots identify the AVFs obtained by incrementally protecting the worst-case component identified through SyRA diagnostic analysis. The green dot represents the best system.

BAYESIAN CROSS-LAYER DSE

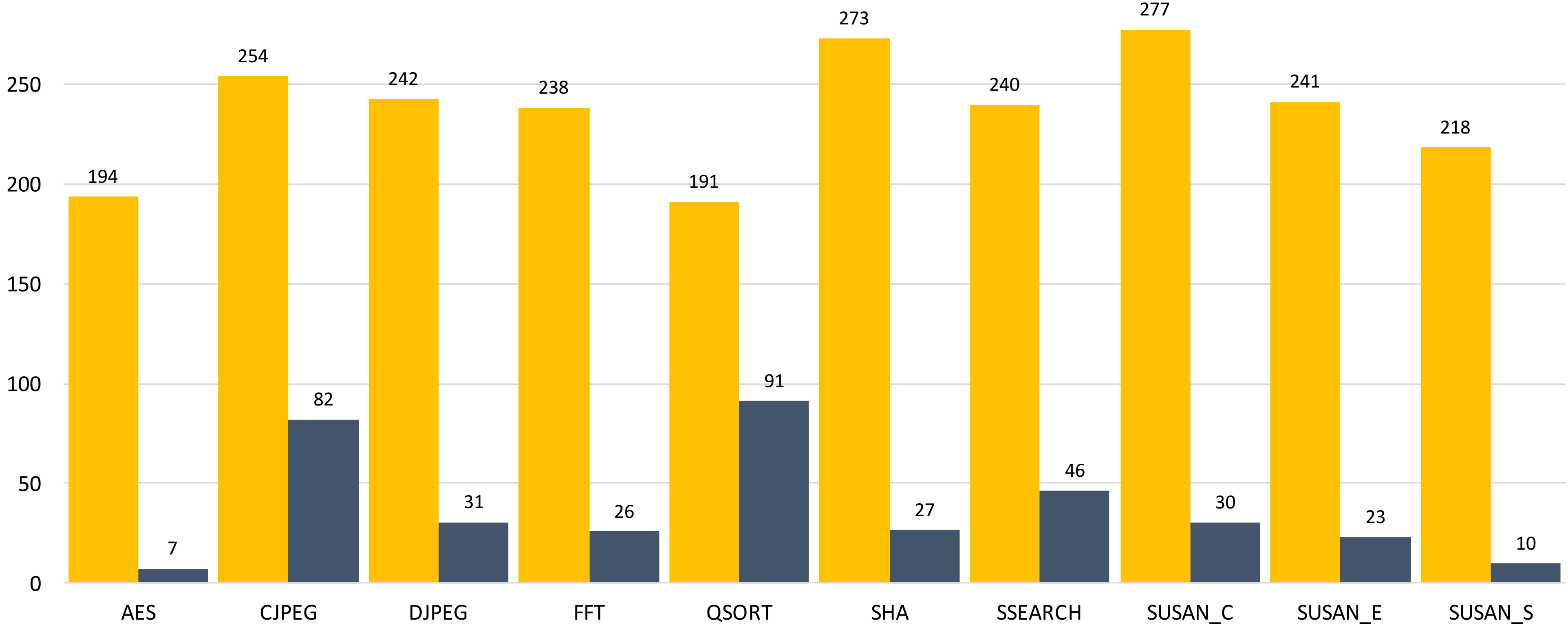
MiBench: AVF improvements



BAYESIAN CROSS-LAYER DSE

MiBench: Number of iterations before best solutions

DSE for best resilience only (yellow bars) vs. DSE for multiple objectives (dark blue bars).



DSE	5.20 5.29	3.45 3.42	3.09 3.24	7.48 8.34	3.47 3.71	6.89 6.98	3.26 3.35	7.38 7.27	8.45 7.36	7.92 5.58
Model Generation	43.90	42.16	14.77	80.39	12.53	16.04	2.06	7.65	11.72	32.75

hours required to generate the model of the initial solution (and all alternative clusters) and to perform the DSE process.

OUTLINE

1. Stochastic Models
2. Bayesian Cross-Layer Analysis
3. Bayesian Cross-Layer DSE
4. Conclusions and perspectives

CONCLUSIONS

Closing remarks and future perspectives

- Resiliency is a critical vector for the whole compute continuum
- Look for solutions that synergize with other design dimensions to reduce the resiliency "tax"
- Better industry support for cross layer resiliency
  - Open standards for HW/SW tunable reliability management structures
  - EDA tools for fast early reliability analysis
  - Challenging from a business perspective but there is precedent for success here…
  - Fosters sharing and openness in a typically closed area of technology…

CONCLUSIONS

Closing remarks and future perspectives

@Test.Group.Polito

https://www.testgroup.polito.it

@TestGroupPolito

Scan me

Questions?

Διαφορετικά, ας πίουμε τσίπουρο