

Modelo de Predição de Desempenho Integrado à Exploração do Espaço de Projetos

Mateus Tostes dos Santos, Thiago R. de Oliveira, Rhayssa Sonohata, Casio Krebs, Liana Duenha, Ricardo R. dos Santos

Trabalho em andamento no âmbito do Programa de Mestrado em Ciência da Computação - FACOM -UFMS.

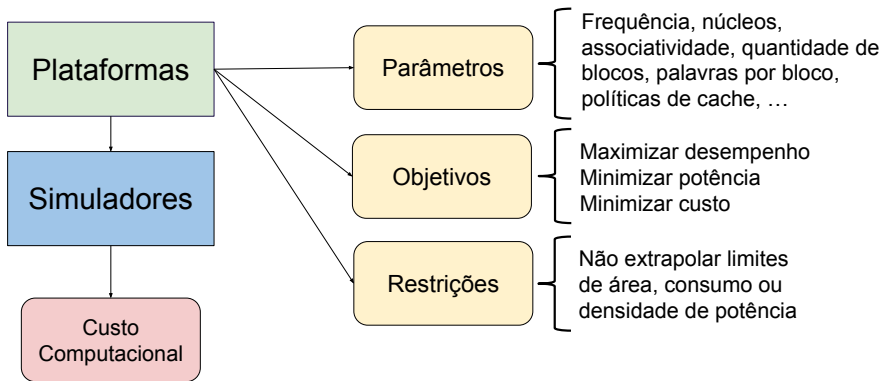
Orientadora: Prof. Dra. Liana Duenha

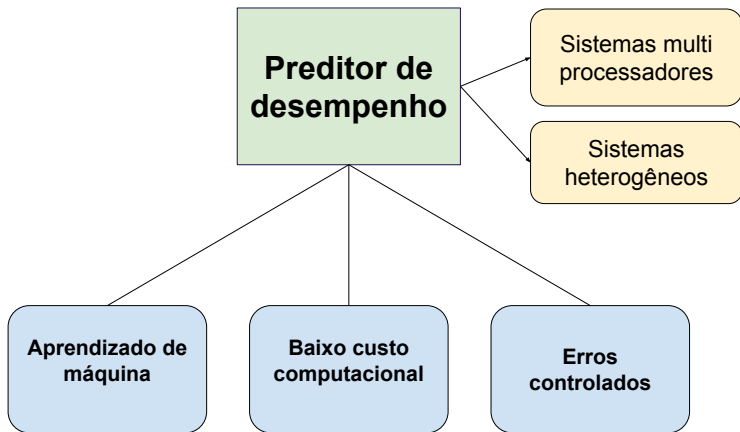
Faculdade de Computação - FACOM - UFMS

October 3, 2018

Motivação

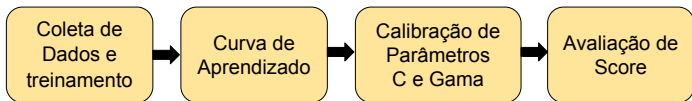
- Plataformas computacionais complexas
- Alto custo computacional para simulação durante o projeto





Preditor de desempenho baseado em SVR

Metodologia para desenvolvimento do preditor de desempenho baseado em SVR



- Kernel RBF ou Gaussiano (função K sobre vetores x_i e x_j do espaço de entradas)

$$K(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$$

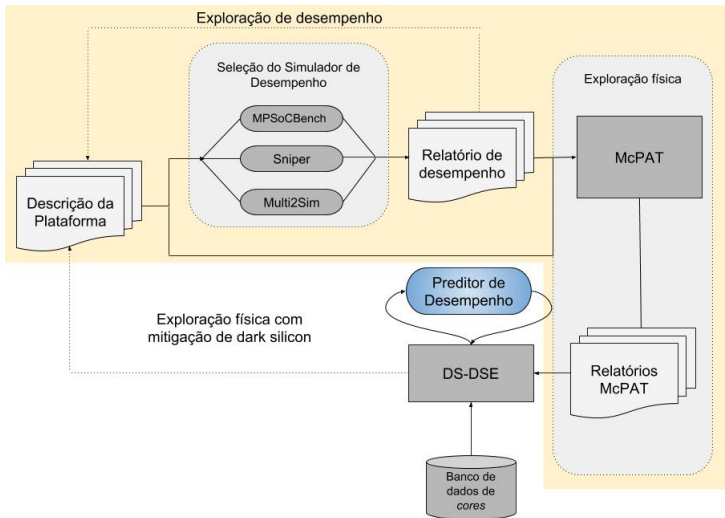
- Avaliação de score:

$$\text{Score} = (1 - u/v)$$

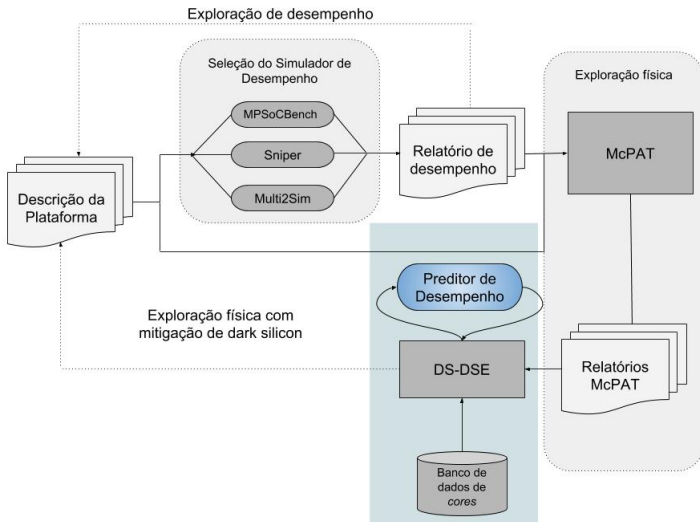
$$u = \frac{\sum_{i=1}^n (\text{resp} - \text{pred})^2}{n}, v = \frac{\sum_{i=1}^n (\text{resp} - \text{media_resp})^2}{n}$$

- **FAFSIM**: Parâmetros arquiteturais e desempenho; sistema *single-core*; **regressão linear**;
- **DVFS+DCT**: OpenMP para verificar o impacto de DVFS e DCT no desempenho e consumo de sistemas *multicore*; **análise estatística**;
- **Turandot**: Predição de desempenho e consumo energético em sistemas *single-core*; **Regressão polinomial com Spline**;

Exploração de espaço de projeto ciente de Dark-silicon (Multiexplorer)



Exploração de espaço de projeto ciente de Dark-silicon (Multiexplorer)



Parâmetros físicos presentes no banco de dados dos núcleos com tecnologia de 32nm

Núcleo	Potência (W)	Área (mm^2)	Desempenho
Smithfield	8,9	9,32	6428
Quark x1000	1,06	6,42	502,53
ARM A53	5,5	7,2	3125,68
ARM A57	12,13	13	4006,64
Atom Silvermont	2,51	5,5	648,47

- Estimativa de desempenho de um core (1/tempo de execução): baixa precisão;
- Estimativa otimista de desempenho de plataformas multicore;

Coleta de dados para treinamento do modelo

- 7.000 simulações realizadas com o SNIPER;
- Estimativa de desempenho considerando heterogeneidade: quantidade de instruções, frequência, quantidade de cores e IPC;
- Modelo de processadores: Smithfield, Atom, Quark, Arm53, Arm57;
- Aplicações: Barnes, FFT, Fluidanimate, Radix, Swaptions;

Calibração do Modelo SVR usando γ , C e Score

	Atom			Quark		
	C	Score	NMSE	C	Score	NMSE
$\gamma = 1$	2500	0.97435	0.00157	2500	0.97571	0.00176
	2000	0.96920	0.00183	2000	0.97117	0.00188
	1500	0.95998	0.00213	1500	0.95970	0.00221
	1000	0.93071	0.00266	1000	0.92216	0.00269
	500	0.72954	0.00496	500	0.70543	0.00562
	100	0.17874	0.10683	100	0.16592	0.11571
$\gamma = 5$	2500	0.97435	0.00130	2500	0.97571	0.00131
	2000	0.96920	0.00132	2000	0.97117	0.00142
	1500	0.95998	0.00171	1500	0.95970	0.00182
	1000	0.93071	0.00247	1000	0.92216	0.00281
	500	0.72954	0.00657	500	0.70543	0.00819
	100	0.17874	0.08557	100	0.16592	0.09404
$\gamma = 10$	2500	0.97435	0.00156	2500	0.97571	0.00170
	2000	0.96920	0.00199	2000	0.97117	0.00222
	1500	0.95998	0.00232	1500	0.95970	0.00258
	1000	0.93071	0.00333	1000	0.92216	0.00378
	500	0.72954	0.01052	500	0.70543	0.01278
	100	0.17874	0.11099	100	0.16592	0.12074

NMSE: erro quadrático médio normalizado;

As 5 melhores soluções encontradas

Cores originais (n_o)	Cores IP (n_{ip})	Tipo do core IP (T_{ip})	Área do chip (A_{mm^2})	Densidade de Potência (dp_t)	Desempenho predito (d_t)	Desempenho original (d_o)
8	12	Quark	151.6	0.55	22086	57454
7	14	Quark	155.12	0.50	19980	52031
6	15	Quark	152.22	0.46	17968	46106
5	17	Quark	155.74	0.40	12871	40683
4	18	Quark	152.84	0.36	11093	34758

- Experimentos utilizando o preditor durante todo o fluxo de DS-DSE
- Explorar Escalabilidade
 - aumentar a quantidade e tipos de cores
 - estender para plataformas many-core (GPUs, FPGAs)
- Explorar Diversidade
 - Novas técnicas de aprendizado de máquina
 - Novos algoritmos de exploração do espaço de projetos

Perguntas?

mateustostesdosantos@gmail.com