

Projeto de GPGPUs Livre de Dark-Silicon com o MultiExplorer

Marco Antônio R. Zárate¹, Laura B. Ludgero¹, Liana D. Duenha¹

¹Faculdade de Computação (FACOM/UFMS)
Campo Grande – MS – Brasil

{marco_zarate, laura_ludgero, liana.duenha}@ufms.br

Abstract. This work presents the enhancement of MultiExplorer as a more robust platform for GP-GPU design-space exploration. Previous usability limitations restricted its analytical potential, but the improvements introduced streamline parameter configuration, heuristic selection, and result interpretation. These refinements strengthen the tool's analytical capabilities, enabling more precise evaluation of energy–performance trade-offs in contemporary architectures constrained by the dark-silicon effect.

Resumo. Este trabalho apresenta o aprimoramento do MultiExplorer como plataforma mais robusta para a investigação do espaço de projetos de arquiteturas GP-GPU. A ferramenta, antes limitada por barreiras de usabilidade, passa a oferecer configuração mais ágil de parâmetros, seleção facilitada de heurísticas e interpretação aprimorada de resultados. As melhorias ampliam seu potencial analítico, permitindo examinar com maior precisão os compromissos entre energia e desempenho em arquiteturas contemporâneas afetadas pelo efeito dark-silicon.

1. Introdução

Durante décadas, a indústria de microeletrônicos foi orientada pela Lei de Moore [Moore 1998]. Contudo, devido a restrições físicas, a redução do tamanho dos componentes eletrônicos torna-se mais complexa. Entre os principais desafios enfrentados estão o aumento do consumo de energia e as limitações térmicas, o que dá origem ao fenômeno conhecido como *dark silicon*.

Este trabalho descreve o aprimoramento do MultiExplorer como uma plataforma mais robusta para investigação sistemática do espaço de projetos de arquiteturas GP-GPU, cujo fluxo é descrito em [Sonohata et al. 2023]. A ferramenta, originalmente concebida para pesquisa acadêmica, apresentava barreiras de usabilidade que limitavam seu potencial para análises exploratórias mais amplas. A evolução proposta neste estudo amplia a capacidade do MultiExplorer de apoiar investigações científicas ao tornar mais eficiente e acessível o processo de configurar parâmetros arquiteturais, selecionar heurísticas de exploração e interpretar conjuntos de resultados complexos.

Este trabalho também se alinha a abordagens semelhantes, como o AccelSim, que realiza simulação validada de GPUs [Yang et al. 2020], e frameworks de DSE com suporte a GUIs, como o COOL [Powell et al. 2019]. Essas iniciativas destacam a importância de ferramentas que unem modelagem quantitativa e usabilidade, objetivo igualmente perseguido pelo MultiExplorer.

2. Aprimoramentos para Exploração do Espaço de Projeto

Interfaces aprimoradas foram desenvolvidas com o objetivo de ampliar a capacidade analítica do MultiExplorer, reduzindo barreiras de entrada associadas ao conhecimento detalhado de seu fluxo interno ou ao uso por linha de comando. Essa evolução torna a ferramenta mais adequada tanto para pesquisadores experientes quanto para novos usuários, fortalecendo sua utilidade como plataforma de investigação em arquitetura de GP-GPUs.

A partir da seleção da arquitetura de referência, da aplicação utilizada para modelagem [Ludgero et al. 2024] e dos valores iniciais de *Streaming Multiprocessors (SMs)*, o usuário pode definir de forma estruturada os parâmetros arquiteturais e os limites físicos que orientarão o processo de exploração. O fluxo subsequente de Exploração do Espaço de Projetos permite avaliar, de maneira sistemática, métricas fundamentais para arquiteturas sujeitas a restrições energéticas e térmicas, incluindo incidência de *dark-silicon*, área, densidade de potência e desempenho.

Além dos resultados de estimativas físicas e de desempenho do sistema original de entrada, o sistema permite realizar a exploração em busca de arquiteturas heterogêneas livres de *dark-silicon* e com a menor degradação de desempenho possível. A exploração arquitetural pode ser com base em uma busca exaustiva ou com base no algoritmo genético aproximado NSGA-II.

Os resultados são apresentados em formato de tabelas e gráficos, que favorecem análise comparativa e interpretação quantitativa, ampliando a capacidade do MultiExplorer de apoiar estudos sobre compromissos energia–desempenho e outras dimensões críticas do projeto de GP-GPUs.

3. Evolução Tecnológica

A evolução tecnológica foi exemplificada através da GPU TITANX, originalmente em 32 nm e sem *dark-silicon*. Reduzindo-se a litografia para 22 nm e aumentando o número de SMs, mantém-se a área e a frequência do chip, porém eleva-se a densidade de potência, o que resulta no surgimento de *dark-silicon*. A Tabela 1 apresenta os valores estimados com base no modelo de cálculo de [Sonohata et al. 2023], indicando cerca de 6,11% de área inativa na configuração da placa em 22 nm.

TITANX (Pascal)		
Número de SMs	24	52
Tecnologia (nm)	32	22
Frequência (MHz)	1000	1000
Área do chip (mm ²)	693,44	693,44
Potência (W)	131,16	139,69
Densidade de potência (W/mm ²)	0,189	0,201
Área de ref. (mm ²)	28,89	13,33
Potência ref. (W)	5,46	2,68
% de <i>dark-silicon</i>	–	6,11%

Tabela 1. Evolução tecnológica da TITANX.

A configuração resultante de 52 SMs em 22 nm é então submetida ao fluxo de DSE, produzindo arquiteturas heterogêneas livres de *dark-silicon*. A Tabela 2 compara

as melhores soluções obtidas pelos métodos exato e heurístico, que apresentam desempenhos semelhantes, com diferença relativa de apenas 0,30%.

Força-bruta	Desemp.	NSGA-II	Desemp.
28T + 11V	1020951	32T + 12V	1017810
31T + 12V	1020543	32T + 11V	1010763
33T + 13V	1018827	32T + 10V	1000000
29T + 11V	1018606	32T + 9V	984894
32T + 12V	1017810	36T + 9V	973802

Tabela 2. Soluções alternativas geradas pelo DSE (SMs em 22 nm).

Observa-se que o NSGA-II encontra soluções próximas ao ótimo com menor custo computacional, embora o método exato permaneça necessário como referência em espaços de projeto maiores.

4. Conclusão

Este artigo apresentou uma versão mais robusta do MultiExplorer, oferecendo suporte mais efetivo à análise do compromisso entre energia e desempenho, que são aspectos centrais na avaliação de arquiteturas contemporâneas, sujeitas às restrições impostas pelo efeito *dark-silicon* dos sistemas desenvolvidos em processos de fabricação modernos.

Para prova de conceito, o artigo mostrou por meio da evolução tecnológica da TITANX, a estimativa de *dark-silicon* prevista na sua evolução tecnológica e ofereceu arquiteturas alternativas heterogêneas livres desse fenômeno. Como trabalhos futuros, prevê-se a inclusão de modelos térmicos, avaliações experimentais mais amplas e suporte a tecnologias de fabricação abaixo de 22 nm. O MultiExplorer está disponível em <https://github.com/lscad-facom-ufms/multiexplorer>.

Referências

- Ludgero, L. B., Damasceno, G. G., Rodrigues, S. S., dos Santos, R. R., and Duenha, L. D. (2024). Exploração do espaço de projetos de sistemas gpppus ciente de dark-silicon. In *Simpósio em Sistemas Computacionais de Alto Desempenho (WIC-SSCAD)*. SBC.
- Moore, G. E. (1998). Cramming more components onto integrated circuits. *Proceedings of the IEEE*, 86(1):82–85.
- Powell, E., Basu, A., and Mishra, P. (2019). Cool: A systematic multi-objective design space exploration framework with graphical interface. In *Proceedings of the 2019 International Conference on Compilers, Architectures and Synthesis for Embedded Systems (CASES)*, pages 1–10.
- Sonohata, R., Arigoni, D. C. A., Fernandes, E. R., dos Santos, R. R., and Duenha, L. D. (2023). Performance predictors for graphics processing units applied to dark-silicon-aware design space exploration. *Concurrency and Computation: Practice and Experience*, 35(17):e6877.
- Yang, X., Kim, S. K., Sengupta, D., Ketterlin, A., Jain, A., and Yalamanchili, S. (2020). Accelsim: An extensible simulation framework for validated gpu modeling. In *2020 IEEE International Symposium on Performance Analysis of Systems and Software (ISPASS)*, pages 53–64.