

Paralelismo SIMD em Modelos Matemáticos SARS-CoV-2

1

Antônio Favarin Freire², Maurício Aronne Pillon³.

¹ Vinculado ao projeto “Impacto energético em nuvens IaaS frente a aplicação de técnicas de escalonamento e de virtualização.”

² Acadêmico (a) do Curso de Bacharelado em Ciência da Computação – CCT – Bolsista PROBIC

³ Orientador, Departamento de Ciência da Computação – CCT – mauricio.pillon@udesc.br

Há mais de um ano, cientistas têm dispensado esforços em pesquisas de áreas distintas do conhecimento com foco no combate a pandemia. Em dezembro de 2019, na época ainda uma epidemia, o SARS-CoV-2, popularmente chamado de coronavírus ou COVID-19, foi descoberto na cidade Wuhan/China. Desde então, estudos sobre tratamento precoce, técnicas inovadoras de concepção de vacinas, tratamentos pós-COVID19, ensaios clínicos de acompanhamento da evolução da doença, técnicas em terapia intensiva são diariamente notícias na mídia. No contexto científico, uma breve pesquisa com a palavra-chave “COVID-19” com filtro de ano a partir de 2020, no *Google Scholar* resultou mais de 179.000 trabalhos.

O SARS (*Severe Acute Respiratory Syndrome* - Síndrome Respiratória Aguda Grave) é uma patologia conhecida da comunidade científica desde 2003. Visto que o COVID-19 é uma doença com alta transmissibilidade, a modelagem matemática é um aliado ao combate da disseminação desenfreada do vírus. Na literatura, as principais abordagens para a modelagem matemática são: *Susceptible-Infected-Removed* (SIR), *Susceptible-Exposed-Infected-Removed* (SEIR), *Susceptible-Infected-Recovered-Dead* (SIRD), *fractional-derivative* SEIR, e SEIRD. Os estados comuns a todas as abordagens são: (S) suscetível, (E) exposto, (I) infectado e (R) removido. A transição de um estado para outro é calculado em função de três taxas: de infecção (β), incubação (λ) e recuperação (γ). O ciclo da doença inicia-se com o (S) – um grupo de pessoas suscetíveis, isto é, não imunizadas e em ambientes propícios a transmissão, é infectada de acordo com uma taxa β de infecção. Uma porção do grupo inicial compõe (E) – grupo de pessoas expostas ao vírus que, em função da taxa λ de incubação gerará o grupo de infectados (I). O ciclo se encerra com (R), removido, o que ocorre em uma taxa γ de recuperação.

No Brasil, a empresa 3778, em parceria com o Hospital Sírio-Libanês, desenvolveu uma ferramenta de apoio ao planejamento de saúde pública (<https://covid-simulator.3778.care/>). O simulador, implementado na linguagem de script *Python* e baseado na biblioteca *Streamlit* (<https://streamlit.io/>), tem acesso on-line e código fonte público (<https://github.com/3778/COVID-19/>). A ferramenta baseia-se no modelo matemático SEIR, oferecendo a implementação *bayes*, *ode* e *sde*. O tempo de resposta da ferramenta on-line é aceitável (segundos) quando a população analisada trata de cidades individualmente. Porém, a ferramenta permite a parametrização e a inclusão de dados de um número maior de cidades, de regiões ou de estados. Em alguns casos, a parametrização gera sobrecarga no processamento e tempos de respostas superiores a minutos.

Este trabalho de iniciação científica analisou o código fonte da ferramenta, mais especificamente, o código legado *seir_bayes*. O foco do trabalho era a análise de viabilidade de paralelismo SIMD para modelos matemáticos SARS-CoV-2. A escolha empírica da ferramenta se deu devido a sua popularidade, disponibilidade do código fonte, e uso do modelo SEIR. O modelo

de arquitetura de processamento SIMD, aplicado a GPU (*Graphics Processing Unit – Unidade de Processamento Gráfico*), trabalha com o conceito de processamento vetorial. A paralelização dos modelos SIR e SEIR em arquiteturas SIMD não é novidade. Na literatura, trabalhos científicos com implementações SEIR em CUDA, OpenMP ou MPI são comuns.

A análise do código sequencial, constituída das etapas de compreensão dos modelos matemáticos, linguagem de script Python e da ferramenta desenvolvida pela empresa 3778, foi concluída com êxito. Na figura 1 demonstra o comportamento do avanço da COVID-19 em uma população aproximada de 13 milhões de pessoas.

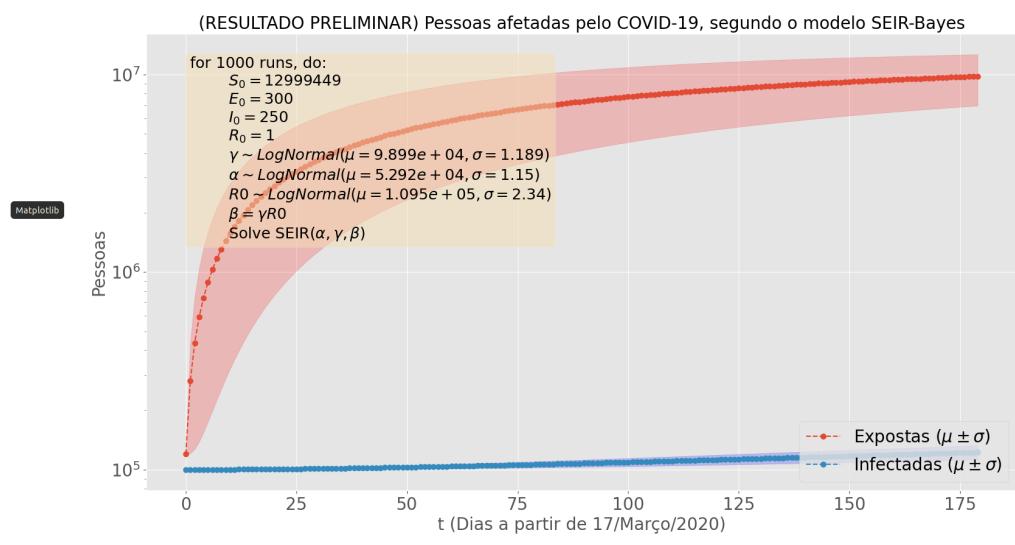


Figura 1. Ferramenta da empresa 3778 com código legado e modelo matemático SEIR-Bayes

Após a verificação dos testes preliminares, a análise do código e os trabalhos científicos estudados permitem concluir que é viável a paralelização SIMD da ferramenta desenvolvida em *Python*. Todavia, o término prematuro do trabalho (meses antes do término do período) não permitiu a incorporação do código paralelo do modelo SEIR-bayes a ferramenta. Como trabalho futuro, tem-se a atualização do código de base do modelo, a escolha do algoritmo paralelo e a implementação em CUDA do algoritmo escolhido.

Palavras-chave: SARS-CoV-2. SIMD. Algoritmos Paralelos.