

IMPACTO DO BUC EM TRAÇOS REAIS ANONIMIZADOS DA GRID'5000

Geremias Corrêa², 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 Ciência da Computação. – CCT – Bolsista PROBIC/CNPq

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

Desde o advento de grades computacionais, há mais de três décadas, os avanços na otimização do fornecimento de recursos apoiaram-se, sobretudo, no monitoramento dos mesmos. Algoritmos de escalonamento tomam suas decisões em função da disponibilidade de recursos físicos e de acordo com as informações das requisições das tarefas. As informações fornecidas pelos usuários permitem estimar a ocupação dos recursos da grade, facilitando a tomada de decisão dos algoritmos de escalonamento. Normalmente, o uso de recursos em grades, tanto comerciais quanto acadêmicas, é contabilizado após o término da execução da tarefa, utilizando a métrica *runtime* (tempo de execução da tarefa) e a quantidade de recursos alocados. Sem penalizações, o *walltime* (tempo máximo estimado de uso) é, quase sempre, ignorada pelo usuário.

Com estimativas imprecisas do uso de recursos, a concepção de previsões baseadas nestas estimativas leva a erros elevados, fazendo com que os algoritmos de escalonamento ignorem esta métrica. A conscientização dos usuários, resultando em estimativas mais precisas, permite a exploração de novos algoritmos de escalonamentos. Este trabalho apoia-se na gamificação, concebendo o algoritmo Beneficiador do Usuário Consciente (BUC), que bonifica usuários preocupados com a precisão de suas estimativas e, pune, aqueles que fornecem, historicamente estimativas imprecisas. O impacto das estimativas é no posicionamento na fila solicitação de recursos. Em casos que a demanda por recursos físicos é maior do que a disponibilidade, tarefas são enfileiradas. O BUC atua bonificando ou punindo os usuários, de acordo com seu comportamento histórico (*score*), no posicionamento na fila.

O funcionamento do BUC é descrito pelas Equações 1 e 2. Na Equação 1, o algoritmo proposto tem uma heurística estática que pondera o tempo de espera atual da tarefa na fila (T), o *walltime* da requisição (W) e o *score* (S), a partir de uma constante K, que determina o peso do *score* com *walltime* da requisição em relação ao tempo de espera, na qual quanto maior o valor R obtido melhor a posição na fila.

$$R = (S * 0,8 + (I/(W/100000)) * 0,2) * K + T \quad (1)$$

O *score* (S) foi uma forma de representar a confiabilidade do usuário, variando entre 0 a 100, em que 100 representa o mais confiável, sendo ela individual a cada usuário e que se atualiza a cada tarefa finalizada, baseado na precisão (P) do tarefa submetida do usuário, como mostrada na Equação 2, na qual é dada pelo tempo real de execução da tarefa (R) dividido pelo *walltime* (W) da tarefa, variando o valor de P entre 0 e 1, em que 1 representa máxima precisão. Execuções em que o tempo real de execução ultrapassam o *walltime* são descartadas.

$$P = R / W \quad (2)$$

Com o intuito de mensurar o impacto no posicionamento inicial da tarefa na fila, utilizou-se traços reais anonimizados da grade computacional acadêmica GRID'5000. O simulador de grades utilizado foi o BatSim (versão disponível em <https://www.pypi.org/project/batsim-py/>). A infraestrutura física e traços utilizados foram do sítio de Lille, referentes ao ano de 2019, compostas por cargas reais (2785) e injetadas (300). A injeção de tarefas busca popular todas as categorias com usuários. Foram realizados testes com duas variações de K para o algoritmo, 1440 e 10800, respectivamente atribuídas ao BUC1440 e o BUC10800.

Os resultados preliminares podem ser observados na Figura 1. A base de comparação foi o FCFS. Os usuários Real 1, 2 e 3 foram escolhidos por dois fatores, maior número de tarefas submetidas no período e a categoria que representam, que são um usuário ruim, médio e bom de forma média, no contexto da plataforma, em que Real 1 tem S médio (2,7), o Real 2, (45,93), e o Real 3, (68,22). Os três usuários Fictícios injetados representam categorias de 0%, 50% e 100% de precisão (P) em todas suas 100 tarefas, de cada, injetadas, sendo seu *walltime* (W) atribuído em 3600 em cada uma delas. Os S resultantes dos usuários Fictício 1, 2 e 3 foram, respectivamente, (4,8), (96,14) e (99).

Os resultados obtidos demonstram a eficiência do algoritmo BUC e permitem a exploração, em trabalhos futuros, de novos algoritmos de escalonamento de tarefas baseados na estimativa de tempo de uso dos recursos. A métrica *Walltime* poderá ser associada a algoritmos de otimização do consumo de energia ou de taxa de utilização de recursos físicos.

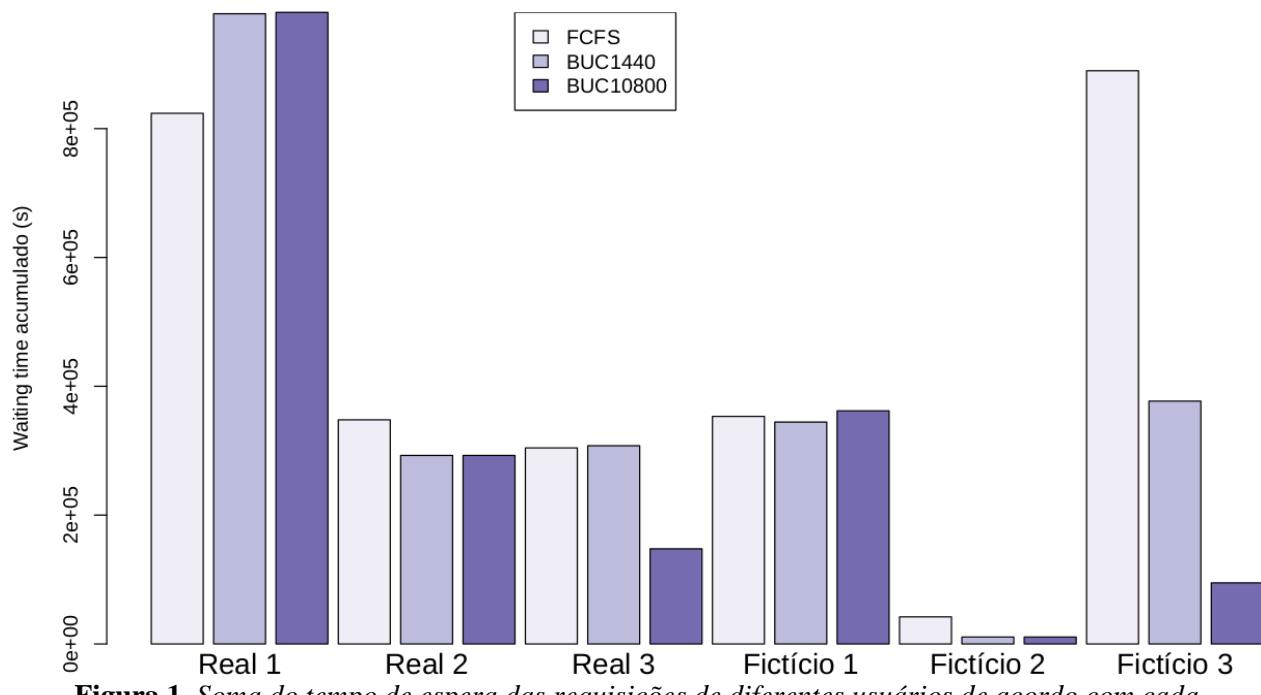


Figura 1. Soma do tempo de espera das requisições de diferentes usuários de acordo com cada algoritmo.

Palavras-chave: Grades computacionais. Algoritmos de escalonamento. Conscientização.