

OpenBatch: Serviço de Processamento em Lote para Nuvem Privada

Marcus Vinicius Paiva Martins¹, Gabriel Felipe Mesquita Esteves¹,
Julia Rodrigues Yoshida¹, Daniel Sundfeld¹

¹Faculdade de Ciências e Tecnologias em Engenharias (FCTE)
Universidade de Brasília (UnB), DF - Brasil

{paiva.marcus, gabriel.esteves, julia.yoshida}@aluno.unb.br

daniel.sundfeld@unb.br

Resumo. *O acesso a ambientes de Computação de Alto Desempenho (HPC) via interfaces de linha de comando impõe barreiras a pesquisadores não especialistas. Este trabalho apresenta o OpenBatch, uma plataforma de código aberto que oferece uma interface gráfica web unificada para submissão, monitoramento e gerenciamento de trabalhos no gerenciador SLURM. A arquitetura desacoplada do sistema compreende um frontend em formato de aplicação de página única (SPA), um backend em Node.js que intermedia operações via SSH e o cluster. A solução destaca-se pelo uso de autenticação delegada ao sistema operacional do nó mestre, garantindo segurança e rastreabilidade sem banco de dados próprio. O OpenBatch fornece múltiplos níveis de abstração, upload de dependências e terminal web integrado via xterm.js, apresentando-se como uma alternativa viável para a otimização de infraestruturas locais existentes sob premissas de nuvem privada e soberania sobre os dados.*

1. Introdução

A crescente demanda por recursos de Computação de Alto Desempenho (HPC) tem impulsionado a pesquisa científica e o processamento de grandes volumes de dados. Tradicionalmente, o acesso a esses ambientes é mediado por interfaces de linha de comando (CLI) [Tanenbaum and Bos 2015], o que exige um conhecimento técnico aprofundado e frequentemente resulta na subutilização do potencial computacional devido à curva de aprendizado íngreme para os pesquisadores.

Embora a computação em nuvem tenha revolucionado esse cenário com serviços de processamento em lote gerenciados (como AWS Batch [Amazon Web Services 2016], Google Cloud Batch [Google Cloud 2026] e Azure Batch [Microsoft 2026]), instituições acadêmicas, governamentais e centros de pesquisa ainda enfrentam desafios na adoção de nuvens públicas. Para essas instituições, a necessidade de previsibilidade de custos, segurança e controle estrito sobre a soberania dos dados torna a infraestrutura de nuvem privada uma alternativa preferencial.

No modelo de nuvem privada, encontramos diversos serviços como projeto de software livre, o OpenStack [OpenStack Foundation 2010] é um dos principais serviços no modelo de infraestrutura como serviço (IaaS) [de Sousa Neto 2015] para gerenciar computação e máquinas virtuais em ambientes locais. Na categoria de serviços

de armazenamento, podemos encontrar exemplos como o TrueNAS (antigo FreeNAS) [iXsystems 2026], um sistema operacional de código aberto usado para transformar computadores comuns em servidores de armazenamento em rede (NAS), permitindo guardar e proteger arquivos de forma centralizada.

No entanto, não conhecemos exemplos que implementem serviços de processamento em lote na nuvem privada. Neste artigo, propomos o OpenBatch: um portal web unificado (Science Gateway) de código aberto que abstrai a complexidade operacional de clusters HPC gerenciados pelo SLURM [Yoo et al. 2003]. O objetivo principal da ferramenta é facilitar o acesso a esses recursos, simplificando o ciclo de vida da computação em lote e oferecendo uma alternativa de baixa complexidade para os pesquisadores continuarem controlando a infraestrutura.

2. OpenBatch

O OpenBatch foi arquitetado seguindo princípios de sistemas distribuídos modernos, adotando um modelo desacoplado em três camadas. A camada de apresentação é composta por uma Aplicação de Página Única (SPA) desenvolvida em HTML, CSS e Javascript. O backend, que atua como a camada lógica da solução, foi implementado em Node.js e expõe uma API RESTful e WebSockets, mediando a comunicação de forma segura por meio de conexões SSH com o nó mestre (plano de controle) do SLURM. A última camada é o próprio cluster HPC, responsável pela execução e armazenamento de dados.

O desenvolvimento priorizou módulos focados na usabilidade e segurança, destacando-se as seguintes implementações: Autenticação Delegada e Segurança, em vez de manter um banco de dados próprio para usuários, o sistema delega a autenticação diretamente ao sistema operacional do nó gerenciador do cluster. Isso garante que os comandos sejam executados sob o contexto de segurança do usuário, permitindo total rastreabilidade (auditoria) sem a necessidade de duplicação de credenciais. A interface possui múltiplos níveis de abstração. Usuários iniciantes contam com um formulário visual simplificado que gera scripts de execução dinamicamente, enquanto usuários avançados dispõem de um editor de texto interativo com realce de sintaxe nativo para submissões customizadas.

A plataforma integra um módulo para upload direto de dependências, códigos e conjuntos de dados empacotados em arquivos ZIP, transferindo-os de forma automática para o diretório local do usuário no cluster. Para facilitar o trabalho, o OpenBatch também oferece um Terminal Web Integrado. Ele foi implementado usando a biblioteca xterm.js e provê um emulador de terminal diretamente no navegador, permitindo a execução de tarefas e simulando fielmente a experiência de uma sessão nativa SSH.

O último módulo, atualmente em implantação, busca a containerização do painel e da submissão de tarefas. Neste modo, todas as camadas do OpenBatch rodam dentro de containers, integrando a execução do SLURM em tarefas disponíveis e executadas dentro de contêineres Docker. A containerização permite uma melhor integração e implantação do OpenBatch com diversas infraestruturas existentes, tornando a execução do OpenBatch independente do sistema operacional disponível.

3. Conclusão e Trabalhos Futuros

A validação funcional do OpenBatch comprova que é viável preencher a lacuna de usabilidade presente em infraestruturas privadas de HPC. O sistema efetivamente centraliza a segurança e oferece múltiplos níveis de abstração, contemplando desde pesquisadores sem especialização em tecnologia até administradores avançados. Ao operar sob um modelo de custo de capital focado na própria infraestrutura (on-premise), o OpenBatch viabiliza um ambiente análogo aos modernos serviços de processamento em lote da nuvem, mas sem as desvantagens do *vendor lock-in* ou dos custos operacionais imprevisíveis. Dessa forma, ele se consolida como uma ferramenta valiosa para instituições que exigem controle soberano de dados e otimização orçamentária no processamento científico e acadêmico. Como trabalhos futuros, desejamos finalizar a implementação de atividades completamente containerizadas. Com o fim desta atividade, poderemos integrar o OpenBatch com nuvens híbridas: produzindo uma fila de trabalho escalável e permitindo que o sistema reconheça picos de demanda e delegue temporariamente o processamento excedente para provedores de nuvem pública. Finalmente, também desejamos melhorar os *dashboards* implementados com uma integração direta com o SLURMdb, a base de dados do SLURM. Também desejamos utilizar algoritmos de análise preditiva no módulo de observabilidade, utilizando dados históricos do cluster para estimar tempos de espera, sugerir partições alternativas e identificar janelas de menor utilização, auxiliando usuários na tomada de decisão.

Referências

- Amazon Web Services (2016). AWS Batch. <https://aws.amazon.com/batch/>. Acesso em: 2026-08-04.
- de Sousa Neto, M. V. (2015). *Computação em Nuvem: Nova Arquitetura de TI*. Brasport.
- Google Cloud (2026). Google cloud batch: A fully managed batch processing service. <https://cloud.google.com/batch>. Acesso em: 2026-08-05.
- iXsystems (2026). TrueNAS: Open Enterprise Storage Platform. <https://www.truenas.com>. Accessed: 2026-08-05.
- Microsoft (2026). Azure Batch Documentation. <https://learn.microsoft.com/en-us/azure/batch/>. Accessed: 2026-08-04.
- OpenStack Foundation (2010). Openstack cloud software. Accessed: 2026-08-04.
- Tanenbaum, A. S. and Bos, H. (2015). *Modern operating systems*. Pearson Education, Inc.
- Yoo, A. B., Jette, M. A., and Grondona, M. (2003). Slurm: Simple linux utility for resource management. In *Workshop on Job Scheduling Strategies for Parallel Processing*, volume 2862 of *Lecture Notes in Computer Science*, pages 44–60. Springer, Berlin, Heidelberg.