



SBGf Conference

18-20 NOV | Rio'25

Sustainable Geophysics at the Service of Society

In a world of energy diversification and social justice

Submission code: WGA97WN4GY

See this and other abstracts on our website: <https://home.sbgf.org.br/Pages/resumos.php>

Cloud execution of Marchenko-based Internal Multiple Modeling using spot instances on a Kubernetes cluster

Alan Albano (Petrobras), Paulo Carvalho (Petrobras), Vitor Batista (Petrobras), Andre Ferreira (petrobras)

Cloud execution of Marchenko-based Internal Multiple Modeling using spot instances on a Kubernetes cluster

Please, do not insert author names in your submission PDF file

Copyright 2025, SBGf - Sociedade Brasileira de Geofísica/Society of Exploration Geophysicist.

This paper was prepared for presentation during the 19th International Congress of the Brazilian Geophysical Society held in Rio de Janeiro, Brazil, 18-20 November 2025. Contents of this paper were reviewed by the Technical Committee of the 19th International Congress of the Brazilian Geophysical Society and do not necessarily represent any position of the SBGf, its officers or members. Electronic reproduction or storage of any part of this paper for commercial purposes without the written consent of the Brazilian Geophysical Society is prohibited.

Introduction.

We describe the challenges and adaptations for a Marchenko-based internal multiple modeling method deployed on a Kubernetes cluster using spot instances from a cloud provider.

The application was developed to model upgoing internal multiples in ocean bottom sensor (OBS) datasets. These datasets are typically organized into common receiver gathers, with receiver gathers designated as the outer loop in the parallelization scheme. This loop could be modeled as a bag of tasks parallelism problem. The application implements intranode parallelism to compute multiples for each trace in the common receiver gather and supports fine-grained checkpointing.

The bag-of-tasks parallelism was implemented using the Ray framework. This framework includes a Kubernetes component called the KubeRay Operator, which provisions and auto-scales Ray clusters using spot instances. Checkpoint data is stored on a managed Lustre filesystem deployed within the cluster.

Theory

The proposed solution minimizes modifications to the existing geophysical application, which currently runs on traditional on-premises HPC clusters. For example, leveraging a managed Lustre filesystem eliminated the need to refactor the application for direct interaction with an object storage service.

Deploying the computation on a Kubernetes cluster proved advantageous, as it enabled leveraging the ecosystem for job orchestration (via Argo Workflows/Events) and the KubeRay Operator, which greatly simplified spot instance utilization. The solution is largely cloud-agnostic, though the managed Lustre filesystem dependency limits portability, as it is not universally available across cloud providers.

Conclusions

The chosen software stack enabled rapid migration of the application to the cloud with spot instance integration. Despite being originally designed for on-premises HPC systems, the geophysical application ran successfully on a Kubernetes cluster using spot instances without requiring code modifications.