Hessisches Kompetenzzentrum
für Hochleistungsrechnen

Parallelizing the Level 2A Processor for Sentinel-2 Satellite Imagery

Project Manager
Taylan Özden

Principal Investigator
Dr. Michael Burger

Project Term
2020 - 2020

Clusters
Lichtenberg Cluster Darmstadt

Additional Software
Sen2Cor, GDAL

Institute
Institute of Scientific Computing

University
Technische Universität Darmstadt

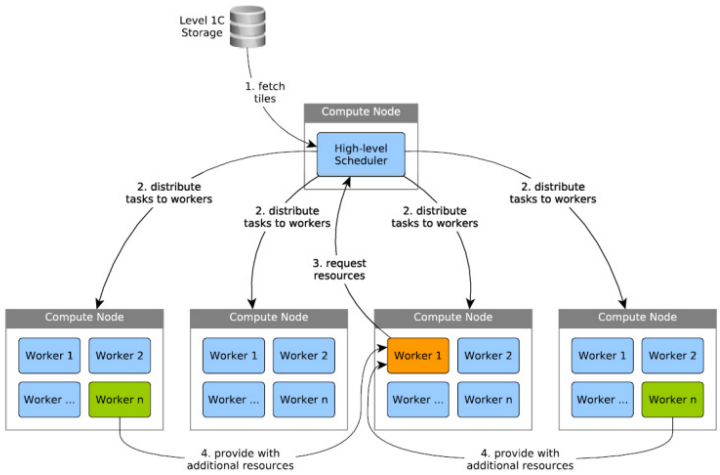

Figure 1: High-level Scheduling Approach.

Introduction

In June 2017, the European Space Agency started systematic Level 2A processing of Sentinel- 2 acquisitions over Europe using the Cloud Screening and Atmospheric Correction Processor named Sen2Cor written in Python. As the Sen2Cor is freely available to the scientific community to allow the processing of Sentinel-2 products with customized configurations, this project develops and tests several strategies to shorten the runtime of the Sen2Cor by employing shared memory approaches. Additionally and in order to facilitate the official deployment of Sen2Cor, this project investigates large-scale scheduling strategies by using modern and platform-independent frameworks to develop an approach for the deployment of Sen2Cor not only on the official ground segment but also on modern systems such as high-performance computers and cloud architectures while still being in line with the current requirements.

Methods

To improve the runtime of the Sen2Cor on shared memory architectures, the multithreaded backend of Joblib and the multiprocessed backend of Dask was employed. The large-scale scheduling approach was established also by using Dask to allow sections which were already parallelized by Dask's backend to be automatically considered in the large-scale scheduler.

Results

JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN
TECHNISCHE UNIVERSITÄT DARMSTADT
UNIKASSEL VERSITÄT
Philipps Universität Marburg

The shared memory approaches developed in the previous project shortened the total runtime of the Sen2Cor by up to 38 % and provided an improved scalability on modern multi-core architectures. The developed large-scale scheduler in this project allows the deployment on the ground segment by fulfilling its requirements and simultaneously providing a scalable and platform-independent system to operate on modern infrastructures.

Discussion

The results demonstrate that parallelizing critical sections by using multithreaded and multiprocessed backends improve the runtime of the Sen2Cor and prove that further extensions can be employed. The large-scale scheduler developed in this project offers a scalable approach to operate on modern infrastructures and allows further integration of tailored scheduling strategies. Additionally, this project demonstrates that unifying both approaches enables an implicit parallelization by using sophisticated scheduling approaches provided by Dask.

Last Update: 2023-03-16 01:30