Additional Application for the Reproducibility Challenge: Experimenting with the Astronomical Image Mosaic Engine, Montage using Datalife

This application offers an opportunity to explore potential performance optimizations for compute-intensive tasks through task parallelization and distribution. We utilize the Montage workflow, originally sourced from Montage Workflow GitHub Repository, with a SLURM script for execution, (See the details https://github.com/pnnl/datalife/tree/main/tutorials/evaluation_scripts/DFL/Montage_DFL).

This page provides instructions for installing Montage, and preparing initial data, so that you can start evaluating performance. Your observations regarding Montage, including task distribution experiments, should be included in the report.

The Montage pipeline runs from the process of downloading astronomical image files (.fits format) across three color bands–red, blue, and green–obtaining a merged single image.

Montage Installation

Download the Montage tarball (v5.0 or higher) from: http://montage.ipac.caltech.edu/

```
tar xvzf Montage_v6.0.tar.gz
cd Montage
mkdir -p bin
make –j
export MONTAGE_PATH=`pwd`/bin
```

Obtain Initial data files from datalife
(https://github.com/pnnl/datalife/tree/main/tutorials/evaluation_scripts/DFL/Montage_DFL)

```
curl –L –o montage_data_16tasks.tar.gz
https://github.com/pnnl/datalife/raw/refs/heads/main/tutorials/evaluation_scripts/DFL/Montage_DFL/montage_data_16tasks.tar.gz
```

A set of files (.tbl and .hdr formats) are prepared to initialize the workflow start.

Deliverables

- Provide a performance plot similar to Figure 6 and provide explanations of potential optimization strategies
- Propose enhancements that balance performance improvements with efficient use of flow resources