**Time-Slice to Time-Series Conversion:**

Modern climate model simulations can produce Petabytes of data. Efficient use of this data in post-processing requires converting it from time-slice files into time-series files. The **PyReshaper** is a parallel Python package for performing this conversion quickly and efficiently.

---

**Parallel Approaches:**

However, serial techniques are still insufficient. There is a tremendous amount of parallelism to exploit in the conversion process, and we investigated two types of parallel techniques for performing the conversion: Task Parallel and Data Parallel.

In the task-parallel approach, each compute rank is responsible for writing one (or more) time-series files. This approach only scales up to the number of variables in the dataset.

In the data-parallel approach, data within each variable is distributed across the compute ranks. This approach can scales well. A prototype tool called the **ncReshaper** was created using this technique.

A hybrid approach is also possible, but it has not been investigated at this point in time.

---

**Inefficient Serial Techniques:**

Current serial techniques, such as the NCO operators, can take an excessively long time with large datasets. Below shows the time to perform the conversion on various CESM datasets. Each dataset spans 10 years and have different grid resolutions. The PyReshaper, in serial mode, performs better on most datasets.

**Serial Duration**

- NCO
- PyReshaper (NetCDF3)

**The Benefits of Parallelism:**

The performance of our new tool can be seen below, compared with other techniques.

- **Parallel Duration**
  - NCO
  - ncReshaper
  - pyReshaper (NetCDF3)
  - pyReshaper (NetCDF4)

- **Parallel Throughput**
  - NCO
  - ncReshaper
  - pyReshaper (NetCDF3)
  - pyReshaper (NetCDF4)

**Availability:**

The **PyReshaper** will be available for download as a stand-alone tool in the very near future. It will be available as part of the CESM workflow in the next version of CESM. For questions about how to obtain and use the PyReshaper, contact Kevin Paul (kpaul@ucar.edu).

---

This work is funded by a National Science Foundation grant, number M0856145.