

Bioinformatics Application Analysis using the PoPLAR Gateway

Nyalia Lui, Yuan Liu, Bhanu Rekepalli

OVERVIEW

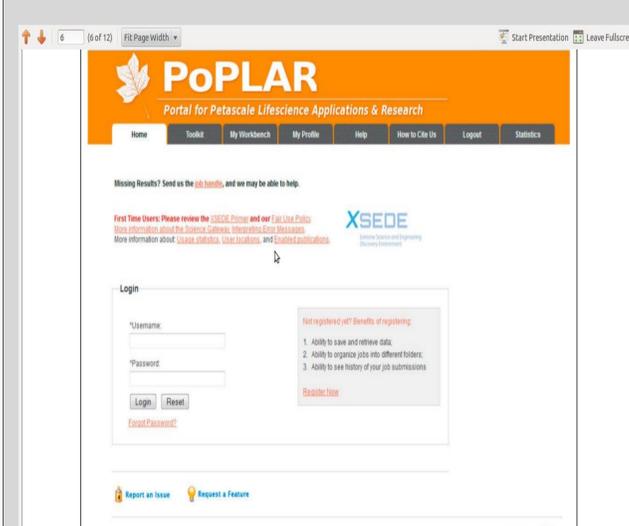
Over the years, scientific data in the life sciences has been generated at exponential rates and continues to grow with the rise of new technologies. We are focusing on the parallel analysis tools and resources used to generate the data. Our focus has led us to using the PoPLAR research project, a scientific gateway designed to provide biologists a means to access High Performance Computing (HPC) resources. Currently, the gateway is under development as a web application with a user friendly interface, eliminating the need for biologists to have basic, and in some cases advanced, computer science skills.

VISUAL OF THE OVERVIEW



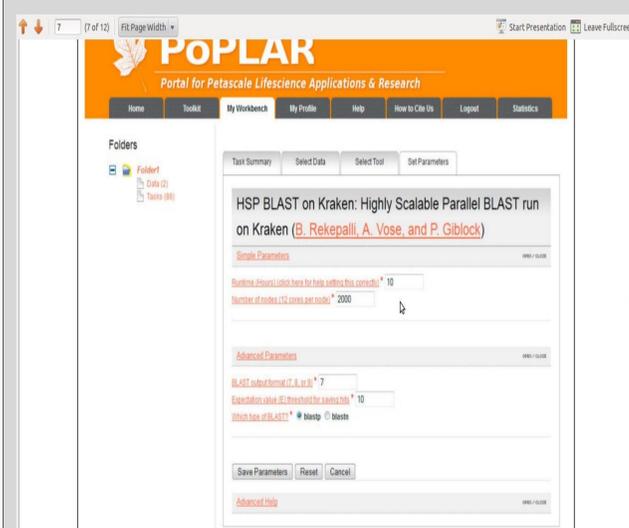
Currently, PoPLAR provides researchers three of the scientific communities most used data analysis tools, BLAST from the National Center for Biotechnology Information (NCBI), HMMER, and MUSCLE.

IMAGES OF POPLAR



Top Image: Is a screen shot of the user login page which will also serve as the home page.

Bottom Image: This is the first screen the researcher will see after logging in. This is where one may submit a new task, check the status of running tasks, or view the results from a finished task.



LIFE OF AN INITIATED TASK

What happens behind the scenes after a researcher has submitted a task to a HPC resource through PoPLAR. The gateway will send all necessary messages to the HPC resource(s) where the desired software is ran. Either the raw output data will be return or the data will go through a back end process in order to parse the data into a specific format.

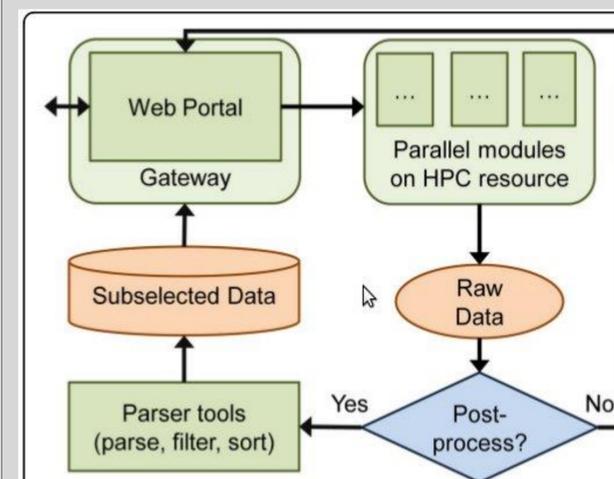
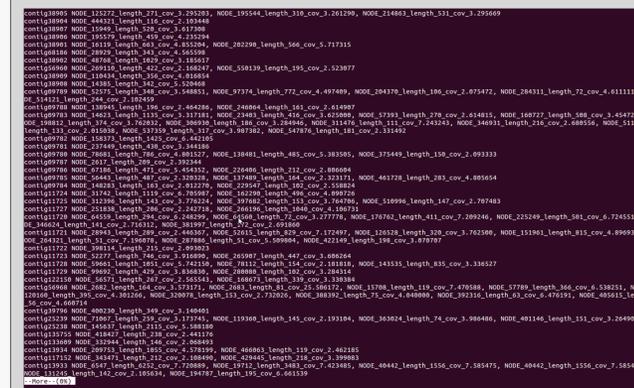


Figure 2 The gateway approach. The user interfaces with the science gateway portal and initiates a job that is run through parallel modules on an HPC resource. Based on the user's needs, either the raw results are delivered back to the user, or the output is parsed and prepared for easier use before it is made accessible through the gateway.

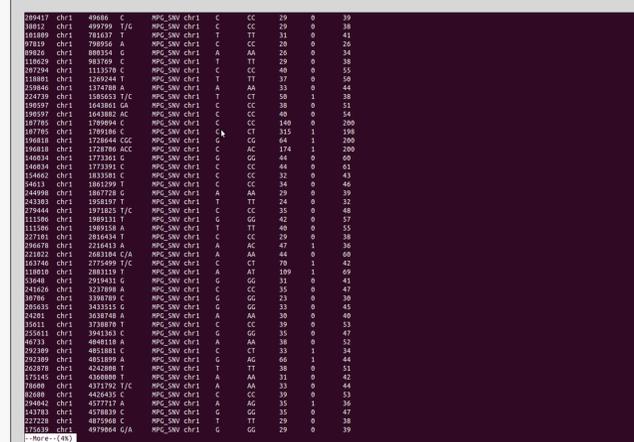
PARSING OUTPUT

To fully understand the workings and different processes that PoPLAR will support, scripts have been written that produce results similar to what a researcher would see when using the gateway. The next column, has the results to two of the scripts.

PARSING OUTPUT IMAGES



Given an output file from a nucleotide BLAST task in tabular format, find all the sequences that did not get hits, and what hits they did or did not get. The image shows the input sequences that got hits and what the hits were.



Given two files the user enters, take Line-X of in one of the files and find matching Line-Y in the other file.

REFERENCES

Rekepalli, Bhanu, Giblock, Paul, Reardon, Christopher. "PoPLAR: Portal for Petascale Lifescience Applications and Research. BioMed Central, Ltd. Web. 17 July. 2013.

Images:
icons.iconarchive.com/icons/aha-soft/people/256/user-icon.png
selab.janelia.org
besocratic.colorado.edu
drive5.com
poplar.nics.utk.edu

CONTACT INFO

Nyalia Lui luinyalia@gmail.com	Yuan Liu yliu89@utk.edu	Bhanu Rekepalli brekepal@utk.edu
--	--	---