Introduction

The National Center for Biotechnology Information (NCBI) offers free access to the Basic Local Alignment Search Tool (BLAST) via the World Wide Web [1][2].

In order to provide a fair and consistent level of service, the NCBI limits the size and number of BLAST searches each user can perform. However, as the cost of sequencing decreases [3], users needing to perform large numbers of BLAST searches may find themselves subject to usage limits. Additionally, users with proprietary or custom data sets are currently forced to use the command line BLAST+ [4] interface.

BLAST @ Amazon Web Services

With the aforementioned users’ challenges in mind we developed the BLAST Amazon Machine Image (AMI) hosted at Amazon Web Services (AWS). This AMI provides users with the ability to instantiate Linux server(s) at AWS which come pre-configured with:

- the latest release of BLAST+
- support for a subset of the NCBI BLAST URL API
- a simplified BLAST web page, and
- A File system in User space (FUSE) client that automatically downloads the most recent copy of popular BLAST databases from NCBI.

System Description

The BLAST AMI is a basic building block to run BLAST on the cloud. It can be used as provided to serve single or multiple users (see Fig. 2) or it can be reused as a modular component in a more elaborate system.

The BLAST AMI acts both as a front- and back-end server. It accepts BLAST searches via its web interface or via a subset of the NCBI BLAST URL API, facilitating integration into existing workflows which use the public NCBI BLAST service.

BLAST searches are executed immediately after submission in the local host, fetching BLAST databases from NCBI as needed and caching them locally. The users of the AMI do not pay for this network traffic.

Additionally, users can log into their instance(s) and use the BLAST+ command line applications or install additional BLAST databases.

Future work

- Explore distribution and partitioning of BLAST searches among multiple nodes
- Provide a cluster-in-a-box solution to run BLAST NCBI welcomes feedback and ideas on how to improve on this experiment.

For more information, visit the NCBI BLAST help page using the following QR code:

Acknowledgements

The authors thank Alexey Iskhakov for the FUSE client, Irena Zaretskaya for web development, Eugene Yaschenko, Don Preuss, and Jim Ostell for helpful discussions and guidance.

References