CloudBioLinux for Data Analysis

- CloudBioLinux is an open-source framework for creating fully automated installation mechanisms for bioinformatics software and data.
- We have contributed numerous extensions and enhancements to CloudBioLinux to make a great environment for building mass spec data analysis platforms.
- With emerging proteomics applications such as proteogenomics, it is becoming essential to build applications platforms that tie together traditional proteomic analyses with other bioinformatic analyses (such as sequence similarity analysis or genomic mapping). CloudBioLinux is an ideal platform for creating such platforms.
- The core CloudBioLinux framework is developed and maintained by Brad Chapman, more information can be found at http://cloudbiolinux.org/.

Applications

- Large Proteomic Tool Suites
  - Trans proteomic-pipeline, OpenMS, crux
- Database Search Tools
  - Myrinmatch, X! Tandem, OMSSA, ...
- Specialized Identification Tools
  - TagRecon, PepNovo, ...
- Validation Tools
  - Percalator, Fido, Mayu, ...
- Bioinformatics
  - NCBI Blast+, EMBASS, Augustus, ...
- Tools Developed at the University of Minnesota
  - iQuant (isobaric quantification), pam-eval (flexible re-evaluation of identified peptide-spectrum matches), peptide-to-gff (map peptides for genomic visualization)

Programming Libraries

pyteomics
- Pyteomics provides a growing set of modules to facilitate the most common tasks in proteomics data analysis…
- https://pypi.python.org/pypi/pyteomics

miprie
- Miprie is a full featured library for working with mass spectrometry data, particularly proteomic, metabolomic and lipidomic data sets...
- https://github.com/pricelab/miprie

R
- Our preconfigured CloudBioLinux image comes preinstalled with various useful R libraries proteomic data analysis (e.g. xcms, mzf, FuctoMineR, caret, ggplot2, VennDiagram, …)

Graphical Applications

The example CloudBioLinux environment comes bundled with user-friendly graphical Desktop applications as well, including:
- MZmine
- PeptideShaker and SearchGUI
- TOPPAS
- PRIDE Converter
- PRIDEInspector

Check out your own copy of CloudBioLinux on Github at https://github.com/chapmanb/cloudbiolinux and build a customized flavor for your mass spec data analysis platform.

- Simple YAML data files describe what software, libraries, and data to install.
- Integrate your own applications (big or small) using native OS packages, fabric functions, Puppet modules, or Chef cookbooks.

CloudBioLinux for Platform Development

- Wine is a compatibility layer allowing many Windows applications to run other operating systems (such as Linux).
- We have built a high level framework for packaging and redistributing Wine environments.
- Includes documentation for creating such a Wine environment and configuring msconvert to work with vendor libraries.
- CloudBioLinux includes support for distributing such environments and creating friendly wrapper scripts along with install procedures for proteomics programs such as mconvert and multiplexr.

CloudBioLinux as a Platform - wine

- Galaxy-P (http://getgalaxyp.org) is an extension to the popular Galaxy (http://galaxyproject.org) framework to enable proteomics workflows.
- The example image that can be launched with CloudBioLinux comes preconfigured with an a subset of Galaxy-P.
- The multi-omics systems biology nature of Galaxy-P demonstrates the utility of building on CloudBioLinux and the complexity of Galaxy-P demonstrates the necessity for automating such deployments.
- Our internal and publicly accessible Galaxy-P servers (https://usegalaxyp.org) are both hosted on a private OpenStack cloud infrastructure and running images built with CloudBioLinux.

CloudBioLinux as a Platform - Galaxy-P

- Swift (https://github.com/romanzenka/swift) is a web interface and workflow engine capable of running Mascot, Sequest, X!Tandem, Myrinmatch, OMSSA, Scafold and IDPicker – extra utilities like mmsvm-air. It provides an unified interface for submitting searches through all the engines simultaneously.
- The CloudBioLinux version of Swift is fully open-source, utilizing only the free tools like mmsvm, X!Tandem, Myrixmatch and IDPicker. This enables Swift to run on as many machines as possible, using Sun Grid Engine for job submission.

CloudBioLinux as a Platform - Morpheus

- Morpheus is a compatibility layer allowing many Windows applications to run other operating systems (such as Linux).
- We have built a high level framework for packaging and redistributing Wine environments.
- Includes documentation for creating such a Wine environment and configuring msconvert to work with vendor libraries.
- CloudBioLinux includes support for distributing such environments and creating friendly wrapper scripts along with install procedures for proteomics programs such as mconvert and multiplexr.

CloudBioLinux as a Platform - Swift

- Swift (https://github.com/romanzenka/swift) is a web interface and workflow engine capable of running Mascot, Sequest, X!Tandem, Myrinmatch, OMSSA, Scafold and IDPicker – extra utilities like mmsvm-air. It provides an unified interface for submitting searches through all the engines simultaneously.
- The CloudBioLinux version of Swift is fully open-source, utilizing only the free tools like mmsvm, X!Tandem, Myrixmatch and IDPicker. This enables Swift to run on as many machines as possible, using Sun Grid Engine for job submission.

Check out your own copy of CloudBioLinux on Github at https://github.com/chapmanb/cloudbiolinux and build a customized flavor for your mass spec data analysis platform.

- Simple YAML data files describe what software, libraries, and data to install.
- Integrate your own applications (big or small) using native OS packages, fabric functions, Puppet modules, or Chef cookbooks.

CloudBioLinux for Platform Development

- Wine is a compatibility layer allowing many Windows applications to run other operating systems (such as Linux).
- We have built a high level framework for packaging and redistributing Wine environments.
- Includes documentation for creating such a Wine environment and configuring msconvert to work with vendor libraries.
- CloudBioLinux includes support for distributing such environments and creating friendly wrapper scripts along with install procedures for proteomics programs such as mconvert and multiplexr.

CloudBioLinux as a Platform - wine

- Galaxy-P (http://getgalaxyp.org) is an extension to the popular Galaxy (http://galaxyproject.org) framework to enable proteomics workflows.
- The example image that can be launched with CloudBioLinux comes preconfigured with an a subset of Galaxy-P.
- The multi-omics systems biology nature of Galaxy-P demonstrates the utility of building on CloudBioLinux and the complexity of Galaxy-P demonstrates the necessity for automating such deployments.
- Our internal and publicly accessible Galaxy-P servers (https://usegalaxyp.org) are both hosted on a private OpenStack cloud infrastructure and running images built with CloudBioLinux.

CloudBioLinux as a Platform - Galaxy-P

- Swift (https://github.com/romanzenka/swift) is a web interface and workflow engine capable of running Mascot, Sequest, X!Tandem, Myrinmatch, OMSSA, Scafold and IDPicker – extra utilities like mmsvm-air. It provides an unified interface for submitting searches through all the engines simultaneously.
- The CloudBioLinux version of Swift is fully open-source, utilizing only the free tools like mmsvm, X!Tandem, Myrixmatch and IDPicker. This enables Swift to run on as many machines as possible, using Sun Grid Engine for job submission.