

Quick and Effective Side-by-Side Installation of Linux on a Windows PC

Mac computers have a Linux terminal built in and can connect with Amazon cloud instances immediately. PC users have a number of options, but creating an Ubuntu Linux volume within the Windows partition is a quick and easy way to install an OS that is compatible with scripts and software packages. This allows many of the scripts you create to also be run locally with no modifications.

Ubuntu Linux can be installed to drive space within the Windows partition without requiring a partition be formatted for it. Many commercial laptops have all possible partitions in use by backup systems which may cause problems if removed. While the space becomes unavailable to Windows, Ubuntu can access all Windows folders and extra drive space outside of the reserved volume. For this reason it is preferable to reserve enough memory for the installation itself and a small active workspace while storing data in Windows folders.

1. Download [Ubuntu LiveCD](#) from ubuntu.com and burn it to a disc or memory stick.
2. Try out the OS by following the LiveCD instructions to boot into trial session. If your computer doesn't boot from CD or memory stick, enter the BIOS and change the boot order from HD > CD > USB to CD > USB > HD.
3. To continue with the installation determine the volume of space to be reserved for Ubuntu. 15-17G is recommended, but 30GB will provide you with an ample workspace for large sequence files. Remember you can access Windows folders and storage from within Ubuntu.
4. Use a defragment/disc cleanup tool. The installer requires contiguous disk space. Files scattered across your hard drive must be moved to create 20-30G of space.
5. Download and run the installer. It will prompt you for the installation size and ask you to for a name and password for the installation account.
6. If the maximum installation size is too small and you have enough free disk space, download a more sophisticated drive cleanup tool.
7. Reboot and select "Ubuntu" from the Windows Boot Manager. Don't walk away during the reboot, you have 10 seconds before Windows automatically loads.

Supplemental Setup Information – Useful Tools for Your Workspace

- "Ubuntu Software Center" on the left sidebar (the grocery bag) has a large number of free programs which can be useful for bioinformatics.
 - KATE – Great text editor for scripting
 - CodeLite – C++ development tool
 - NCBI toolkit – Search for NCBI to find a number of tools
 - Python interpreters and development tools
- BioPython – Set of very useful Python tools with excellent tutorials which can be used to convert, parse, etc., sequences available at biopython.org
- R – Useful for statistics and generating histograms large amounts of data, available at cran.r-project.org
- NCBI standalone BLAST for Linux – Useful on cloud instances, but can be run locally. Available from NCBI homepage under Downloads

Resources

SEQanswers Wiki and Forum

The wiki describes the function of many useful tools and links to the files. There are also great how-to articles which describe some basic workflows and provide reviews for various software packages. The forum is a great resource for advice, help, and scripts from some very talented members.

BioPython Tutorial and Cookbook

Python is a very useful language for parsing sequence data. BioPython has many scripts useful for bioinformatics as built-in functions allowing for easy sequence conversion, generating histograms and statistics, etc. It also has an installation guide.

Python Documentation - docs.python.org/index.html

Contains tutorials and a database of every aspect of Python and sample usages. A bit technical for a quick-reference, but very complete.

LinuxCommand - linuxcommand.org/writing_shell_scripts.php

Beginners guide to shell scripts. How to create workflows and pipelines that will automate otherwise time-consuming tasks and using powerful tools to process large quantities of data quickly.

Software Carpentry – software-carpentry.org

Beginners guides to various programming languages and shell scripts.

Stack Overflow – stackoverflow.com

A question and answer site for any programming language. Posting a question will usually result in a quick reply or link. Very useful for cleaning up scripts and problem solving. Many beginner questions have already been asked and answered with useful snippets of code or scripts.

SeqHack – seqhack.blogspot.com

Scripts related to my bioinformatic projects will be archived at this website.