DNALinux virtual desktop edition

Article *in* Nature Precedings · August 2007

DOI: 10.1038/npre.2007.670.1 · Source: OAI

2 authors:

Sebastian Bassi
National University of Quilmes
10 PUBLICATIONS **190** CITATIONS
SEE PROFILE

Virginia González
Toyoko LLC
10 PUBLICATIONS **91** CITATIONS
SEE PROFILE

DNALinux: Virtual Desktop Edition (VDE)

Abstract

The new version of DNALinux (VDE) is presented. DNALinux VDE is a departure from traditional distributions since it uses a virtual machine to bundle together the operating system and bioinformatics applications. The main advantage of this approach is that a virtualized environment doesn't affect a installed system. With a virtual machine a Linux system can be run under a Windows system, provided that the virtual machine player is installed. The included programs are listed and specifications to add more programs are explained. We believe that DNALinux could be used as a standardized virtual machine for learning, using, developing and testing bioinformatics applications.

Introduction

The Universidad Nacional de Quilmes and GenesDigitales released the first version of DNALinux on January 2004. It was called BioShell (version 0.12) at that time, but it rapidly changed the name to a DNALinux (version 0.13).

That version was based in Slax Linux [1], that was a customizable live CD. A live CD is a CD that can be used to boot the computer and it runs the Operation System and applications from the CD (hence the "live" term) without the need to install in the hard disk. The main advantage of this approach is that the user doesn't need to modify any existing installation. The drawback is the relative low speed of loading applications, high memory usage and it can't run side by side with current operating system. This version of DNALinux was a CD with an Operating System and bioinformatics applications ready to use. The impact of this distribution was better than expected since it was downloaded more than 4000 times from more than 100 countries. There were four versions released as live CD (from 2004 to 2007).

In recent years there was a noteworthy advance in virtualization technologies, so we decided to explore this path. For this reason we changed the type of our Linux distribution. Instead of using a live CD, we built a virtual machine and we called it "DNALinux Virtual Desktop Edition".

A virtual machine is a software that creates a virtualized environment between the computer platform and its operating system, so that the end user can operate software on an abstract machine [2]. This way a Linux operating system can be executed on a virtual machine that is installed in Windows. This version is an image of a virtual machine with preinstalled bioinformatics applications. This time DNALinux is based on Xubuntu, the light version (adapted for low end machines) of the popular Ubuntu Linux. Xubuntu is shipped with top quality packages from the Ubuntu archives but is faster than Ubuntu because it uses the Xfce Desktop environment [3]. The choosen virtual machine is VMWare [4], that provides a free virtual machine player that runs under both Windows and Linux operating system.

Bioinformatics software included in DNALinux VDE

1. NCBI BLAST 2.2.16
2. NCBI NetBLAST 2.2.16
3. EMBOSS 4.1.0
4. ESIM4 (EMBOSS SIM4)
5. MSE 1.0.0

6. Phylip 3.6b
7. Biopython 1.43
8. BioPerl
9. Kalign 1.04
10. Clustalx and Clustalw 1.83
11. njplot
12. GNU Polyxmass 0.9.7
13. Primer3
14. Rasmol 2.7.2.1.1
15. Sigma-align 1.0
16. t_coffe 2.50
17. xviewg
18. Treeviewx 0.5.1
19. hmmer 2.3.2
20. NCBI-epcr
21. Cn3D NCBI Database Viewer
22. DDV Sequence Alignment Viewer
23. Entrez NCBI Database Querying Tool
24. OneD Biological Sequence Viewer
25. PyMOL Molecular Graphics System
26. Sequin DNA Sequence Submission Tool

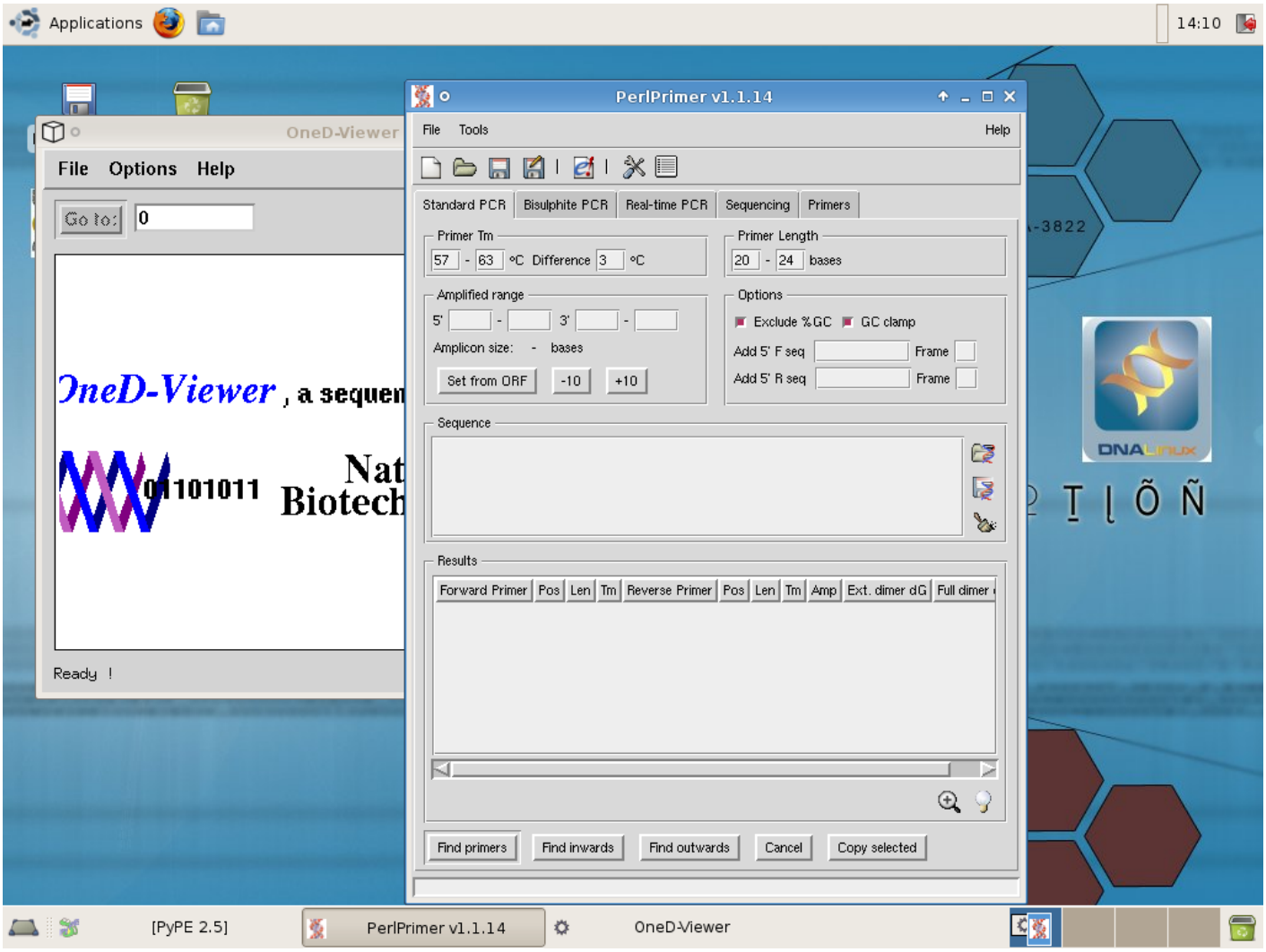
Criteria for inclusion of new software into DNALinux

We are eager to include more software into DNALinux. There are technical and legal specifications that must be met to include software in DNALinux distribution. From a technical perspective the software should be able to be installed in a 32 bits x86 Linux (Kernel 2.6). The disk size of the virtual machine is limited to 20Gb, so applications that have massive storage requirement are discouraged. From a legal point of view, the license must make the program free to distribute. Any license approved by OSI or FSF will do it. If your code have a proprietary license not listed there, please make sure that the license allows us to freely distribute the software.

The future

The end user experience is what will drive future changes into DNALinux. To gather user experience we need to have more users downloading DNALinux VDE. In this moment we are looking for web or ftp servers to host the virtual image (with high bandwidth since the image size is more than 900Mb). Quilmes University datacenter can't handle all the load. Linux distributions are being consolidating into Ubuntu so a good advice to bioinformatics developers is to try target their code to this platform. Since not all developers can change their operating system, a virtual machine could be used as a learning or testing environment.

Screenshot



Resources

DNALinux webpage: http://www.dnalinux.com
Announcements Mailing List: http://groups.google.com/group/dnalinux/
Xubuntu webpage: http://www.xubuntu.org/
VMWare Player: http://vmware.com/download/player/
Bioinformatics Organization: http://www.bioinformatics.org
Bionformatics mailing list: http://groups.google.com/group/general-bioinformatics
Bionformatics mailing list (Spanish): http://groups.google.com/group/bioinformatica-es

References

[1] Tomas Matejicek. Slax. Available at: http://www.slax.org/. Accessed August 9, 2007.
[2] Wikipedia contributors. Virtual machine. Wikipedia, The Free Encyclopedia. August 7, 2007, 04:22 UTC. Available at: http://en.wikipedia.org/w/index.php?title=Virtual_machine&oldid=149696304. Accessed August 9, 2007.
[3] Olivier Fourdan. Available at: http://www.xfce.org/. Accessed August 9, 2007.
[4] VMware. Available at: http://www.vmware.com. Accessed August 9, 2007.