Linux operating system
libre, free: how and why

Dobrica Pavlinušić

http://www.rot13.org/~dpavlin/

PBF, 27.03.2008.

History
Open Source
GPL license

Don't panic!

- What is operating system?
- Linux
- Open Source
- Evolution of open source software
- How is all of this connected with bioinformatics?

Operating system?

- An operating system (OS) is a set of computer **programs** that **manage** the **hardware** and **software** resources of a computer
- The operating system forms a **platform for** other system software and for application **software.**

Software
Applications
Operating system
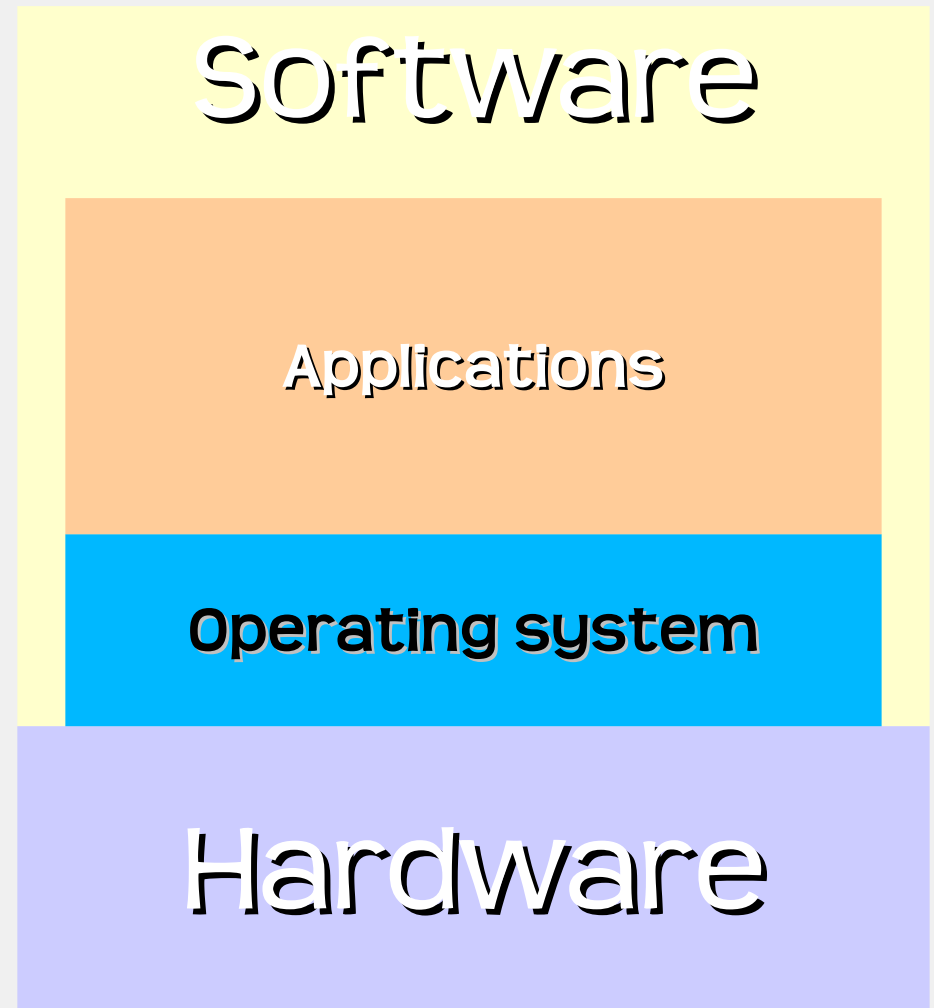
Hardware

Linux FUD*

- hard to install, support for different hardware
  - try recent distribution (eg. Ubuntu)
- it's ugly
  - Gnome, KDE, XFCE ... (are they all ugly?)
- something free can't be good
  - ?
- it works different than Windows
  - some would call this benefit :-)

Linux history from 1991.

- Linus Torvalds wanted OS compatible with Unix for his i386 PC
- Unix is based on POSIX standards
- rapid releases on Internet under GPL license
- result is POSIX complaint OS

If you want to travel around the world and be invited to speak at a lot of different places, just write a Unix operating system. --Linus

Linux and GNU project

- kernel on it's own isn't enough
- user-land already existed as part of GNU project
- Every compilation on Linux system is done by GPL software written as part of "Gnu is Not Unix" project

We all know Linux is great. It does infinite loops in 5 seconds. --Linus

GNU project

- 1983
  - Richard M. Stallman begins development of free Unix compatible operating system
- 1990
  - most of the components (libraries, compliers, shells) are finished
  - development of GNU Hurd kernel starts (still pending)

GPL license and source code

GPL is viral: sharing of changes in GPL software is requirement to the benefit of community

Basically, I want people to know that when they use binary-only modules, it's THEIR problem. I want people to know that in their bones, and I want it shouted out from the rooftops. I want people to wake up in a cold sweat every once in a while if they use binary-only modules. --Linus

4 GPL freedoms

0: to run the program, for any purpose

1: to study how the program works, and adapt it to your needs (source code)

2: to redistribute copies so you can help your neighbor

3: to improve the program, and release your improvements to the public, so that the whole community benefits (source code)

Open Source evolution

- Let's take program as an unit of evol.,
- Every change (improvement, adaptation) is evolution step
- If program is not open (source, specification) it will die out with next change of OS or hardware platform (environment change)
- Every `vi` implementation today descends from Bill Joy's original `vi`
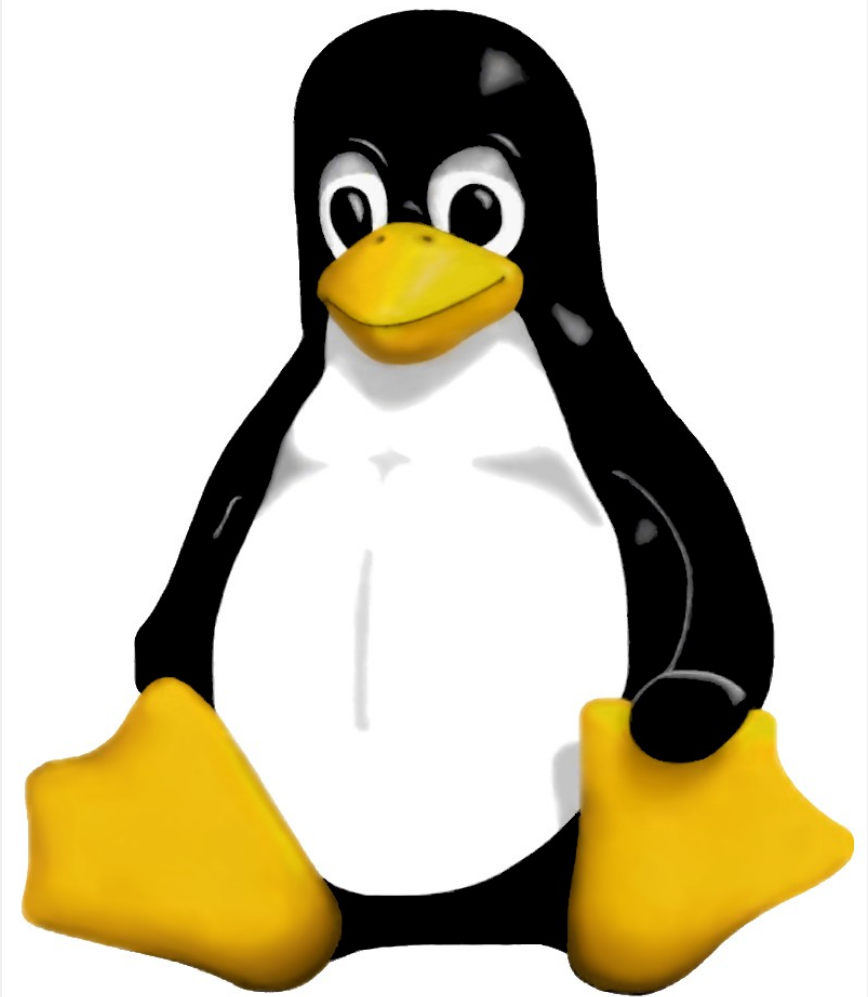- Think of all programs for VAX or DOS

Linux

Software

- Applications
- GNU user-land
- Linux kernel

Hardware

Kernel versions

- history
  - stable (1.0, 1.2, 2.0*, 2.2, 2.4, 2.6)
  - development (0.9, 1.1, 1.3, 2.1, 2.3, 2.5)
- 2.6.X – stable (in distributions)
- 2.6.X-rcY – development (git)

*) released in 1996. with 64 bit support for Alpha CPU

'Fewer fundamental changes' is a mark of a system that isn't evolving as quickly, and that is reaching middle age. We are probably not quite there yet --Linus

Device support

- PDA
- smart phones (GSM+PDA)
- ADSL modems
- set-top boxes (digital TV)
- computers: servers, laptops, clusters, super-computers...

- supports more than 20 CPU architectures

GNU/Linux distributions

free and libre

- Fedora
  - rpm
- Debian, Ubuntu
  - deb
- Gentoo
  - portage

commercial with support

- RedHat Advanced Server
- SuSE/Novell

- Linux Standard Base

Applications

- solving new problems require innovative solutions
- balance between learning and customizations (power users)

- desktop
  - office tasks (OpenOffice.org), web (Firefox), video (VLC), pictures (GIMP)...
- servers
  - web servers, web services (Google)

Programming languages

Ada awk Brainf*ck Basic C C++
Erlang Forth GNAT Haskell
Java KL1 KMFL Lisp Logo Lua
ML Oberon-2 OCaml Perl PHP
Pike PostScript Prolog Python
Ruby Scala Scheme sh SPL
TCL ...

Software development

- Cathedral
  - source available with each release
  - developed by thight group of developers

- Bazaar
  - public view of source on Internet
  - rapid quick releases

Linus' law: given enough eyeballs, all bugs are shallow

Bioinformatics

- Information technologies used for biology
- **Lots of data**, lots of challenges

- different programming languages
- specialized libraries
  - BioPerl
  - BioJava
  - BioBike (LISP)

- access to data through Web services

Computational biology

- using simulations to check assumptions
- computer as a tool

- NCBI provides DNA sequence search using web service
- BLAST – well known search algorithm
- with simple perl or python script you can start right away!

Open standards

- Operating systems based on POSIX – enables compilation of same code on different architectures
- Exchange of data
  - formats (ASCII, XML, JSON, …)
  - protocols (HTTP, SOAP, REST, …)

Open Source culture

- ides are too universal to be used only in software development
  - free culture
  - science
  - education
  - politics
- Creative Commons
  - set of licenses for content

Predators

What are threats to our ecosystem?

Digital Rights Management

- Digital **Restrictions** Management
- DRM enable **chain of trust** between **hardware** manufacturer, **software** supplier and **content** provider
- **User** is left out of this equation (no control whatsoever), so he can just **rent** services, not own anything (can't even change hardware configuration – like bolted car hub)

Patents

- doesn't protect small inventor
- slows down software development
- mostly used as deterrent against other patents
- problem of innovative **ideas** in society where everybody **builds on work of others** (prior art)
- limited in length (hence, slowdown as opposed to standstill with copyright)

Monoculture

Why is diversity necessary?

How is Linux different?

- Requires more up-front knowledge
- Most of actions are direct (by commands), often in command line
- Separation of tasks into small reusable tools, suitable for scripting
- Total control (for customizations), limited only by user's knowledge
- Ability to re-use other people knowledge down to source code

Where is the money?

- Commodity infrastructure
  - OS, web server, database
- Innovative services (Google)
- Quick to innovate
  - alpha geeks

I'm never in the situation where I have to make a priority-decision between Linux and money - all my Linux-related work can be done purely on technical issues rather than having any "marketing" issues pop up. --Linus

Review

- Linux kernel
- GNU project
- GPL license (4 freedoms)
- Distributions
- Customizations
- Bioinformatics, computational biology
- DRM, patents

Software

Bioinformatics

Applications

GNU user-land

Linux kernel

Hardware
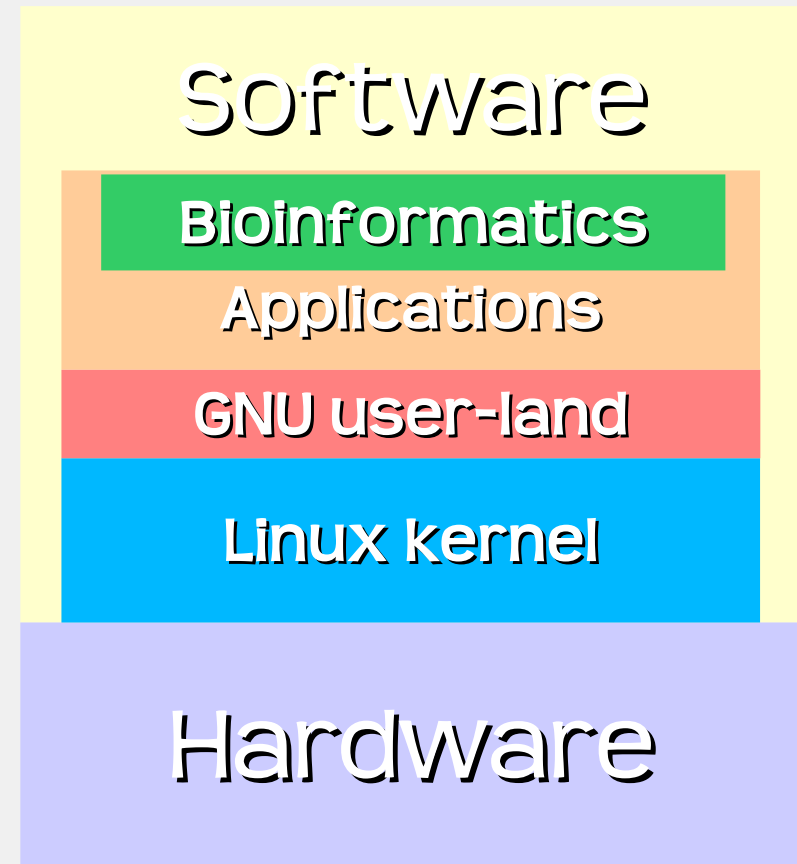
cooperation is only way to **progress** and **sharing** is only way of cooperation

Find out more...

- E.S. Raymond: **"The Cathedral and the Bazaar"**, http://www.catb.org/~esr/writings/cathedral-bazaar/
- Pekka Himanen: **"The Hacker Ethic and the Spirit of the Information Age"**, ISBN 0-375-50566-0
- Feller, B. Fitzgerald, S. A. Hissam and K. R. Lakhani: **"Perspectives on Free and Open Source Software"**, MIT Press, http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=11
- D.E. Geer, C.P Pfleeger, B. Schneier, J.S. Quarterman, P. Metzger, R. Bace, P Gutmann: **"Cyberinsecurity: The Cost of Monopoly --How the Dominance of Microsoft's Products Poses a Risk to Security,"** Computer and Communications Industry Association, September 24, 2003: http://www.ccianet.org/papers/cyberinsecurity.pdf

Questions?
42