

Gentoo Prefix as a Physics Software Manager

Benda Xu (Tsinghua University), Guilherme Amadio (CERN),
Fabian Groffen (Gentoo), Michael Haubenwallner (Gentoo)

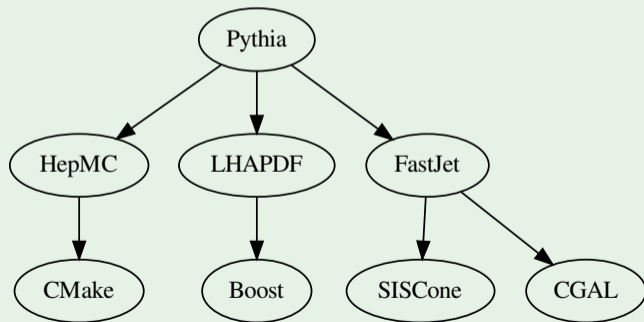
Gentoo Linux
<https://www.gentoo.org>

Nov. 5, 2019

Sophisticated Analysis and Simulation

- Data analysis and detector simulation advance fast.
- Physics software stack is getting deeper.

Example (Pythia physics event generator)



Brace the Complexity

- Bundling dependencies all together is not elegant.
 - ▶ Modularity is crucial for maintainability.
- Package manager is the industry standard to solve this.
 - ▶ It is invented by GNU/Linux distributions.
 - ▶ Debian founder Ian Murdock (1973–2015):
the single biggest advancement Linux has brought to the industry.

Brace the Complexity

- Bundling dependencies all together is not elegant.
 - ▶ Modularity is crucial for maintainability.
- Package manager is the industry standard to solve this.
 - ▶ It is invented by GNU/Linux distributions.
 - ▶ Debian founder Ian Murdock (1973–2015):
the single biggest advancement Linux has brought to the industry.
- Every GNU/Linux distribution has its package manager.

Example (Install the Z Shell)

```
apt-get install zsh      #Debian
dnf install zsh         #Redhat
pacman -S zsh           #Archlinux
zypper install zsh     #SUSE
emerge zsh              #Gentoo
```

Characteristics of Gentoo

- Full GNU/Linux distribution (2000–)
 - ▶ General purpose: from daily to scientific use cases.
 - ▶ Portability: x86, amd64, sparc, arm/64, alpha, mips, riscv, ...
 - ▶ Meta-distribution: ultimate flexibility for specific needs.

Characteristics of Gentoo

- Full GNU/Linux distribution (2000–)
 - ▶ General purpose: from daily to scientific use cases.
 - ▶ Portability: x86, amd64, sparc, arm/64, alpha, mips, riscv, ...
 - ▶ Meta-distribution: ultimate flexibility for specific needs.
- Community driven: spontaneous and distributed development.
 - ▶ Packages, like the build recipe of `Pythia`, are shared pieces of wisdom of the community.

Characteristics of Gentoo

- Full GNU/Linux distribution (2000–)
 - ▶ General purpose: from daily to scientific use cases.
 - ▶ Portability: x86, amd64, sparc, arm/64, alpha, mips, riscv, ...
 - ▶ Meta-distribution: ultimate flexibility for specific needs.
- Community driven: spontaneous and distributed development.
 - ▶ Packages, like the build recipe of Pythia, are shared pieces of wisdom of the community.

Example (`emerge -t pythia`)

```
[ebuild N ] sci-physics/pythia-8.2.26:8::gentoo
[ebuild N ] sci-physics/hepmc-2.06.09-r1::gentoo
[ebuild N ] sci-physics/lhapdf-6.2.3::gentoo
[ebuild N ] dev-libs/boost-1.71.0:0/1.71.0::gentoo
[ebuild U ] dev-util/boost-build-1.71.0::gentoo
[ebuild N ] sci-physics/fastjet-3.0.6-r1::gentoo
[ebuild N ] sci-physics/siscone-3.0.3::gentoo
```

Gentoo Package Manager for Physics

Gentoo Prefix installs a complete Gentoo userspace into a directory, called `${EPREFIX}`.

```
$ tree -L 1 -d
```

```
HOME/gentoo
```

```
|-- bin  
|-- etc  
|-- lib  
|-- lib64  
|-- run  
|-- sbin  
|-- tmp  
|-- usr  
`-- var
```

- ELF Program Headers INTERP points to dynamic loader in `${EPREFIX}`.

Gentoo Package Manager for Physics

Gentoo Prefix installs a complete Gentoo userspace into a directory, called `${EPREFIX}`.

```
$ tree -L 1 -d
```

```
HOME/gentoo
```

```
|-- bin  
|-- etc  
|-- lib  
|-- lib64  
|-- run  
|-- sbin  
|-- tmp  
|-- usr  
`-- var
```

- ELF Program Headers INTERP points to dynamic loader in `${EPREFIX}`.
- dynamic loader reads `${EPREFIX}/etc/ld.so.conf` and load libraries from `${EPREFIX}/lib`, etc.

Gentoo Package Manager for Physics

Gentoo Prefix installs a complete Gentoo userspace into a directory, called `${EPREFIX}`.

```
$ tree -L 1 -d
```

```
HOME/gentoo
```

```
|-- bin
|-- etc
|-- lib
|-- lib64
|-- run
|-- sbin
|-- tmp
|-- usr
`-- var
```

- ELF Program Headers INTERP points to dynamic loader in `${EPREFIX}`.
- dynamic loader reads `${EPREFIX}/etc/ld.so.conf` and load libraries from `${EPREFIX}/lib`, etc.
- Gentoo Prefix is a self-contained userspace,
 - ▶ and share the global filesystem view.

Inside Gentoo Prefix

- autotools-based build systems

```
./configure --prefix="${EPREFIX}" ...
```

Inside Gentoo Prefix

- autotools-based build systems

```
./configure --prefix="${EPREFIX}" ...
```

- CMake-based build systems

```
cmake -DPREFIX="${EPREFIX}" ...
```

Inside Gentoo Prefix

- autotools-based build systems

```
./configure --prefix="${EPREFIX}" ...
```

- CMake-based build systems

```
cmake -DPREFIX="${EPREFIX}" ...
```

- Python, Perl, R, Haskell packages

inherit language interpreter directory prefix.

Inside Gentoo Prefix

- autotools-based build systems

```
./configure --prefix="${EPREFIX}" ...
```

- CMake-based build systems

```
cmake -DPREFIX="${EPREFIX}" ...
```

- Python, Perl, R, Haskell packages

inherit language interpreter directory prefix.

- Toolchain

`gcc`, `binutils` `sysroot=${EPREFIX}`, search for headers and libraries in `EPREFIX`.

`gcc` inject `EPREFIX` dynamic linker.

`glibc` look for configurations in `EPREFIX/etc`.

Reference: <https://goo.gl/wQEkhE>

Interoperability with Other Package Managers

Thanks to the flexibility of portage, other package formats can be straightforwardly converted into ebuilds. A few examples:

[R CRAN](#) Gentoo R-Overlay

[Python PyPI](#) Gentoo PyPI ebuild generator

[Emacs ELPA](#) Gentoo ELPA ebuild generator

[Octave Forge](#) Gentoo Octave overlay

[Java Maven](#) Gentoo Maven overlay (on-going)

[Rust Carge](#) Gentoo Cargo ebuild generator (on-going)

[Julia Pkg](#) Gentoo Julia overlay...

[Conda Forge](#) Gentoo Conda ebuild generator...

[SpackDev](#) Gentoo Spack ebuild generator...

Interoperability with Other Package Managers

Thanks to the flexibility of portage, other package formats can be straightforwardly converted into ebuilds. A few examples:

R CRAN Gentoo R-Overlay

Python PyPI Gentoo PyPI ebuild generator

Emacs ELPA Gentoo ELPA ebuild generator

Octave Forge Gentoo Octave overlay

Java Maven Gentoo Maven overlay (on-going)

Rust Cargo Gentoo Cargo ebuild generator (on-going)

Non-existent, you are welcomed to try :)

Julia Pkg Gentoo Julia overlay...

Conda Forge Gentoo Conda ebuild generator...

SpackDev Gentoo Spack ebuild generator...

Packages in Gentoo

- Almost 20000 official packages.

```
/var/db/pkg $ find . -maxdepth 2 -and -type d | wc -l  
19732
```

```
/var/db/pkg $ ls -l | grep ^sci
```

```
sci-astronomy
```

```
sci-biology
```

```
sci-calculators
```

```
sci-chemistry
```

```
sci-electronics
```

```
sci-geosciences
```

```
sci-mathematics
```

```
sci-physics
```

```
sci-visualization
```

- R Overlay: 18993 packages, e.g. emerge `ggplot2`
- You can do anything on supercomputers:
 - ▶ Install the latest KDE (K Desktop Environment) suite.
 - ▶ Install your own `slurm` by Gentoo to distribute jobs.

Packages in Gentoo

- Almost 20000 official packages.

```
/var/db/pkg $ find . -maxdepth 2 -and -type d | wc -l  
19732
```

```
/var/db/pkg $ ls -1 | grep ^sci
```

```
sci-astronomy
```

```
sci-biology
```

```
sci-calculators
```

```
sci-chemistry
```

```
sci-electronics
```

```
sci-geosciences
```

```
sci-mathematics
```

```
sci-physics
```

```
sci-visualization
```

- R Overlay: 18993 packages, e.g. emerge `ggplot2`
- You can do anything on supercomputers:
 - ▶ Install the latest KDE (K Desktop Environment) suite.
 - ▶ Install your own `slurm` by Gentoo to distribute jobs.

Our manifesto

LIVE on the supercomputer!

Case Study 1: a Stone-Aged Computing Cluster

- Default environment is RHEL (Redhat Enterprise Linux) 5.

RHEL	Release	End of Security Updates	Lastest Version
5	2007-03-15	2017-03-31	5.11
6	2010-11-10	2020-11-30	6.10
7	2014-06-10	2024-06-30	7.7
8	2019-05-07	2029-05-??	8.0

Case Study 1: a Stone-Aged Computing Cluster

- Default environment is RHEL (Redhat Enterprise Linux) 5.

RHEL	Release	End of Security Updates	Lastest Version
5	2007-03-15	2017-03-31	5.11
6	2010-11-10	2020-11-30	6.10
7	2014-06-10	2024-06-30	7.7
8	2019-05-07	2029-05-??	8.0

```
$ uname -a
Linux ln0 2.6.18-194.17.1.0.1.el5_lustre.1.8.5 #1 SMP \
  Wed Aug 3 18:09:09 CST 2011 x86_64 x86_64 x86_64 GNU/Linux
$ cat /etc/redhat-release
Red Hat Enterprise Linux Server release 5.5 (Tikanga)
$ python -c "import sys; print sys.version"
2.4.3 (#1, Jun 11 2009, 14:09:37)
[GCC 4.1.2 20080704 (Red Hat 4.1.2-44)]
$ ld --version
GNU ld version 2.17.50.0.6-14.el5 20061020
Copyright 2005 Free Software Foundation, Inc.
$ ldd --version
ldd (GNU libc) 2.5
Copyright (C) 2006 Free Software Foundation, Inc.
```

Case Study 1: Short Cut to Modern Age

- Prefix: no restriction by the host OS:

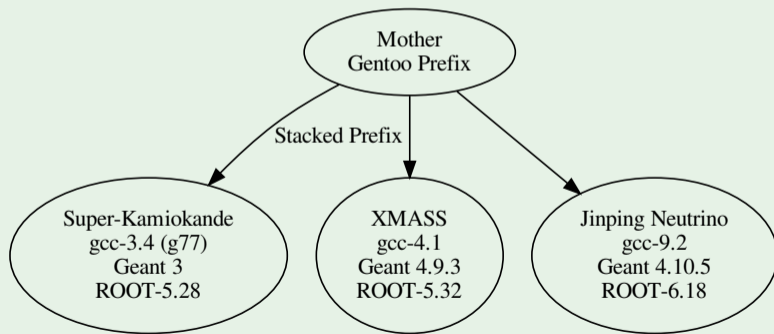
```
$ python -c "import sys;print(sys.version)"
3.6.6 (default, Aug 31 2018, 03:10:49)
[GCC 8.2.0]
$ ld --version
GNU ld (Gentoo 2.30 p1) 2.30.0
Copyright (C) 2018 Free Software Foundation, Inc.
$ ldd --version
ldd (Gentoo 2.19-r1 p3) 2.19
Copyright (C) 2014 Free Software Foundation, Inc.
(glibc-2.19 is the last glibc supporting linux-2.6.18 of RHEL 5)

$ ocaml --version
The OCaml toplevel, version 4.05.0
```

Case Study 2: Simultaneous Multiple Versions

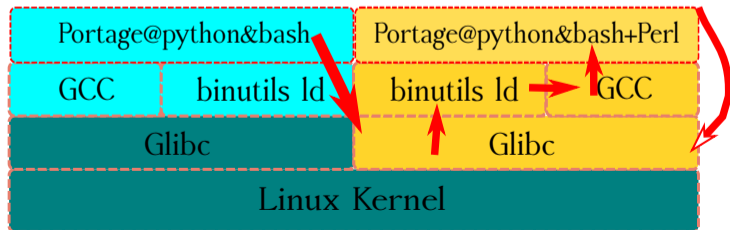
- Stacked-Prefix can make a new Prefix without duplication of disk space.
 - ▶ Coexistence of software suites.

Example (Legacy of experimental physics)



How to Try It Out

- Download and run the bootstrap script:
ref. <https://wiki.gentoo.org/wiki/Project:Prefix>
`wget https://goo.gl/czXbjP -O bs.sh`
`chmod +x bs.sh`
`./bs.sh`
- bootstrap has 3 Stages:
 - 1 Compile python and portage in `${EPREFIX}/tmp`.
 - 2 Install gcc in `${EPREFIX}/tmp` by portage.
 - 3 Build Gentoo in `${EPREFIX}` by portage and Stage 2 gcc.



Android devices

- Gentoo on Android: default EPREFIX=/data/gentoo
Ref.: <http://wiki.gentoo.org/wiki/Project:Android>

- Alternative normal user package managers
 - ▶ Anaconda (specific, 2012–)
 - ▶ Homebrew (general, 2009–)
 - ▶ GNU Guix (general, 2013–)
 - ▶ Spack (specific, 2013–)
 - ▶ Easybuild (specific, 2012–)

Conclusion

- Alternative normal user package managers
 - ▶ Anaconda (specific, 2012–)
 - ▶ Homebrew (general, 2009–)
 - ▶ GNU Guix (general, 2013–)
 - ▶ Spack (specific, 2013–)
 - ▶ Easybuild (specific, 2012–)
- Gentoo (general 2000–) has been the paradise for geek users and developers for decades.
 - ▶ From the community, but share the mindset of scientists.
- The physics use case of Gentoo is a **natural consequence** of its flexibility and expertise in building operating systems.
 - ▶ Free ride on the developments outside science.

G. Amadio and Benda Xu, *Portage: Bringing Hackers' Wisdom to Science*, <https://arxiv.org/abs/1610.02742>