RISC-V support for Gentoo Prefix

Gentoo GSoC 2022 Proposal

Mentors: Guilherme Amadio(Gentoo) amadio@gentoo.org>

Kenneth Hoste(EESSI) < kenneth.hoste@ugent.be>

Applicant: Atharva Amritkar(wiredhikari) < atharvaamritkar@protonmail.com >

Contact Information

Name : Atharva Amritkar

Email : atharvaamritar@protonmail.com

IRC : <u>@atharvaamritkar:matrix.org</u>

GitHub : wiredhikari

Phone : +91 7620261907

Website : <u>atharva.io</u>

Twitter : <u>wiredmuggle</u>

Address: Azad Hall, Indian Institute of Technology Kharagpur, India

Location: Kolkata, India (GMT +5:30)

Age : 20

Project Information

Title : RISC-V support for Gentoo Prefix

Duration: 350 hours

Time : 12 Weeks

Technologies: Bash, Gentoo Linux, RISC-V

Topics : CPU Architecture

Abstract

RISC-V is an emerging open standard instruction set architecture that can be used to build hardware without paying royalties. Gentoo Prefix develops and maintains a way of installing Gentoo systems in a non-standard location, designated by a "prefix". Prefix installs a complete Gentoo userspace into a directory by running the bootstrap-prefix. sh script. The script launches the interactive installer which will ask for some information and discuss some options based on the system being bootstrapped. Prefix has support for several architectures, and we plan to add support for RISC-V. We will make it possible to bootstrap, and use a Gentoo Prefix system on RISC-V architecture.

Objectives

- To make a completely new profile for Prefix to support RISC-V Architecture.
- Make it possible to bootstrap Stage-1, Stage-2 and Stage-3 via an automated bootstrap process for RISC-V machines.
- Provide a working Gentoo Prefix installation on RISC-V architecture.
- Test and keyword packages in Gentoo for RISC-V.
- Thorough documentation to make it easy to add support to upcoming Architectures.

Implementation

I have access to the RISC-V machine (via ssh), so I will be working on that machine or QEMU if needed. The first week is being allocated for creating a Profile for RISC-V architecture. Although, it will get patches throughout the process.

The automated bootstrap script bootstrap-prefix.sh will install the Prefix in three stages: Stage-1 will install tools that are needed by portage to function and then eventually compile python and portage in \$(EPREFIX)/tmp. Stage-2 will start emerging build utilities (one from the base profile) to ensure that dependencies are met. After this, portage will use the dependency system, and all the packages will emerge from the target system. This is necessary for the system to function correctly. In Stage-3, the tree of ebuilds is updated to the latest and flags can be set. Recompilation of all packages installed earlier with custom optimizations and features will take place. We will use GCC installed by portage to install a base system in \$EPREFIX. After these three stages, the prefix will be ready to use and we can install packages via portage. Following that, I will test and keyword the packages.

I will debug and solve the errors as they pop up while getting all the three stages to work and make necessary changes in the bootstrap-prefix.sh script. While working with the script, I will lint the codebase and make it more readable for other contributors. I will write weekly blogs to share my progress and add utility scripts, to reproduce the tasks for transparent evaluation and feedback, which will prevent vagueness in reproducing the work later. Documentation of the process will eventually help in documenting "Porting Prefix to new Architecture" which will help to add support for new architectures in future.

Timeline

Pre-Community Bonding period:

During this period, I have gained access to the RISC-V machine with the help of PLCT-Lab (via ssh). For backup, I have set up Dlan's <u>riscv_image</u> and Ryan's <u>riscv-chroot-env</u>. As everything is already in place, I have started working early by reading wiki articles, and fixing RISC-V and Prefix related bugs. I will discuss with the mentors to improve the current execution plans. I have also set up a personal <u>blogging site</u> for documentation.

Community Bonding Period:

I am familiar with Gentoo systems, so I will tinker with Prefix source code to understand how the project is structured. I will work on structuring and creating the new RISC-V Profile and test the necessary packages. This will help reduce the resistance during the coding period.

Coding Period:

Through the coding period, I will be fixing issues that occur during bootstrapping. I will post updates on my blog and be in communication with my mentors on IRC channels. I shall be available for weekly discussions / meets to discuss my progress and plan the upcoming weeks. So here is the tentative time distribution of the execution plan.

Week 1:

I will start by creating a new Profile, making necessary changes in the bootstrap-prefix.sh script, and ensuring it works for the bootstrapping process for the RISC-V architecture. This shall allow Stage-1 to continue.

Week 2-3:

After having a new profile for RISC-V, we will ensure Stage-1 installs coretools required for Portage to function. Then the latest python and portage will be installed in $\{PREFIX\}/tmp$. We will debug the issues faced during the installation and add patches to the RISC-V profile.

Week 4-5:

After getting a working Stage-1 and making a new profile, Stage-2 will start with emerging build utilities. Portage will use the system's compiler to install GCC into \${EPREFIX}/tmp; then coretools from the base profile. Portage will use the dependency system to emerge system. Eventually, this will lead to using Portage to emerge as a base system.

Week 6-8:

After getting Stage-2 to work, we will use GCC installed by Portage to install a Gentoo base system in \${EPREFIX}. Continue emerging some of the core toolchain packages that make sure we compile and link everything taking the Prefix into account. Glibc is expected to cause trouble; we will fix the issues. Then portage will rebuild all installed packages with custom optimizations and features. Work on LLVM and musl builds.

Week 9-10:

Compile contents in the blogs written so far and document "Porting Prefix to new Architectures". Do rigorous testing and fix bugs, finalize features, and any refactoring (if needed). Take reviews from mentors and other developers and see if any improvements can be made. During this period, we will be able to bootstrap Prefix.

Week 11-12:

These two weeks are left out as a buffer in case of any unforeseen delays or events that will affect the timeline of this project. If everything goes well and there are no delays, I would be more than happy to use this time to add more support and fix bugs. If time allows and everything goes as planned, I will start working on the Post GSoC ideas I have mentioned below.

Deliverables

- A working Profile for RISC-V architecture.
- Bootstrap and use a Gentoo Prefix system on RISC-V architecture.
- Test and keyword necessary packages in Gentoo for RISC-V.
- Documentation on "Porting Prefix to new Architectures".
- Thorough documentation of all work done via blogs and Gentoo Wiki.

Post GSoC

- Adding <u>NixOS support</u> to Gentoo Prefix.
- Provide RISC-V support for Prefix if any bugs occur in the future.
- Continue to maintain and keyword packages for RISC-V.
- Contribute to <u>pkgcraft</u> and <u>cargo-ebuild</u>.
- Maintain and update documentation.

Commitment to Gentoo

I have always had an interest in exploring and working with new technologies, especially Linux and Systems Programming. Gentoo is my daily driver and with my effort, I can make it even better for other users. I have been trying to contribute to the Gentoo Community for several months now. A mentorship program like Google Summer of Code will provide me with great exposure to working with the Gentoo team and help me become a long-time contributor. I want to be a part of the community irrespective of selection in GSoC.

I have my final exams which will be over by the mid of April, and following that, I have a 4-month long holiday. I will be able to devote 40+ hours a week, and my work timings are pretty flexible as well. I have submitted no other proposal for Google Summer of Code. If I get selected, I shall be available during the entire coding period to work on the project.

Below are some of my contributions to Gentoo:

- Bugzilla:
 - o <u>#835069</u>
 - o <u>#786813</u>
- Testing Packages for RISC-V keywords:
 - o <u>#24904</u>
 - o #24630
 - o <u>#837806</u>
- Personal Overlay:
 - hikari-overlay
- Others:
 - NixOS support for Prefix
 - Pkgcraft

Biography

I am Atharva Amritkar, a sophomore pursuing a Bachelor's at the Indian Institute of Technology, Kharagpur. I have been using Gentoo as a daily driver for two years now. I have prior experience in DevOps and had done an internship in the past, which involved working with Gentoo-based servers. I was a Community Edition developer for EndeavourOS in 2021. I have been a mentor in the open-source winter of codes of GWoC and KWoC. Apart from that, I lurk around the Unofficial Gentoo Discord Server and Gentoo IRC channels. In my free time, I rice my desktop and play chess. I program in Rust, Bash, Python, and C. I have been working with RISC-V recently. You can find my work on GitHub.

References

Adding ppc64le support to Prefix:

https://bugs.gentoo.org/755551

RISC-V QEMU image:

https://dev.gentoo.org/~dlan/riscv/

RISC-V Chroot environment:

https://gitlab.com/cwittlut/riscv-chroot-env

Gentoo Prefix as Physics Software Manager:

http://dx.doi.org/10.1051/epjconf/202024505036

Portage: Bringing Hackers' Wisdom to Science:

https://arxiv.org/pdf/1610.02742.pdf

Maintaining Large Software Stacks in a Cray Ecosystem with Gentoo Portage: https://cug.org/proceedings/cug2016 proceedings/includes/files/pap122s2-file1.pdf

Gentoo Wiki for Porting:

https://wiki.gentoo.org/wiki/Porting

Gentoo Wiki for Profiles:

https://wiki.gentoo.org/wiki/Profile_(Portage)

Gentoo Wiki for Prefix:

https://wiki.gentoo.org/wiki/Project:Prefix