Processs-in-Process (PiP)

2.0.0

Refernce Manual

December 10, 2020

Generated by Doxygen 1.8.5

Contents

Chapter 1

Proces-in-Process (PiP) Overview

Process-in-Process (PiP)

PiP is a user-level library to have the best of the both worlds of multi-process and multi-thread parallel execution models. PiP allows a process to create sub-processes into the same virtual address space where the parent process runs. The parent process and sub-processes share the same address space, however, each process has its own variable set. So, each process runs independently from the other process. If some or all processes agree, then data owned by a process can be accessed by the other processes. Those processes share the same address space, just like pthreads, but each process has its own variables like the process execution model. Hereinafter, the parent process is called PiP process and sub-processes are called PiP tasks.

PiP Versions

Currently there are three PiP library versions:

- Version 1 - Deprecated
- Version 2 - Stable version
- Version 3 - Stable version supporting BLT and ULP (experimental)

Unfortunately each version has unique ABI and there is no ABI compatibility among them. The functionality of PiP-v1 is almost the same with PiP-v2, however, PiP-v2's API is a subset of the PiP-v3's API. Hereafter **NN** denotes the PiP version number.

Bi-Level Thread (BLT, from v3)

PiP also provides new thread implementation named "Bi-Level Thread (BLT)", again, to take the best of two worlds, Kernel-Level Thread (KLT) and User-Level Thread (ULT) here. A BLT is a PiP task. When a PiP task is created it runs as a KLT. At any point the KLT can becomme a ULT by decoupling the associated kernel thread from the KLT. The decoupled kernel thread becommes idle. Later, the ULT can become KLT again by coupling with the kernel thread.

User-Level Process (ULP, from v3)

As described, PiP allows PiP tasks to share the same virtual address space. This mans that a PiP task can context-switch to the other PiP task at user-level. This is called User-Level Process where processes may be derived from the same program or different programs. Threads basically share most of the kernel resources, such as address space, file descriptors, a process id, and so on whilst processes do not. Every process has its ows file descriptor

space, for example. When a ULP is scheduled by a KLT having PID 1000, then the getpid() is called by the U-LP returns 1000. Further, when the ULT is migrated to be scheduled by the other KLT, then the returned PID is different. So, when implemnting a ULP system, this systemcall consistency must be preserved. In ULP on PiP, the consistency can be maintained by utilizing the above BLT mechanism. When a ULT tries to call a system call, it is coupled with its kernel thread which was created at the beginning as a KLT. It should be note that Thread Local Storage (TLS) regions are also switched when switching ULP (and BLT) contexts.

Execution Mode

There are several PiP implementation modes which can be selected at the runtime. These implementations can be categorized into two;

- Process and
- (P)Thread.

In the pthread mode, although each PiP task has its own static variables unlike thread, PiP task behaves more like PThread, having a TID, having the same file descriptor space, having the same signal delivery semantics as Pthread does, and so on. In the process mode, a PiP task behaves more like a process, having a PID, having an independent file descriptor space, having the same signal delivery semantics as Linux process does, and so on. The above mentioned ULP can only work with the process mode.

When the **PIP_MODE** environment variable set to "()thread" then the PiP library runs in the pthread mode, and if it is set to "process" then it runs in the process mode. There are also three implementations in the process mode; "process:preload," "process:pipclone" and "process:got." The "process:preload" mode must be with the **LD_PRELOAD** environment variable setting so that the clone() system call wrapper can work with. The "process:pipclone" mode is only effective with the PIP-patched glibc library (see below).

Several function are made available by the PiP library to absorb the functional differences due to the execution modes.

License

This package is licensed under the 2-clause simplified BSD License - see the LICENSE file for details.

Installation

Basically PiP requires the following three software packages;

- `PiP` - Process in Process (this package)
- `PiP-Testsuite` - Testsuite for PiP
- `PiP-glibc` - patched GNU libc for PiP
- `PiP-gdb` - patched gdb to debug PiP root and PiP tasks.

By using PiP-glibc, users can create up to 300 PiP tasks which can be dbugged by using PiP-gdb. In other words, without installing PiP-glibc, users can create up to around 10 PiP tasks (the number depends on the program) and cannot debug by using PiP-gdb. Note that PiP will not run at all without PiP-glibc on CentOS/RedHat 8.

There are several ways to install the PiP packages; Yum (RPM), Docker, Spack, and building from the source code. It is strongly recommended to use the following PiP package installation program (pip-pip):

- `PiP-pip` - PiP package installing program

This is the easiest way to install PiP packages in any form. Here is the example of `pip-pip` usage:

```
$ git clone https://github.com/RIKEN-SysSoft/PiP-pip.git
$ cd PiP-pip
$ ./pip-pip --how=HOW --pip=PIP_VERSION --work=BUILD_DIR --prefix=INSTALL_DIR
```

HOW can be one of `yum`, `docker`, `spack` and `github`, or any combination of them. `pip-pip --help` will show you how to use the program. `yum`, `docker` and `spack` include all three packages; PiP-glibc, PiP-lib, and PiP-gdb.

PiP Documents

The following PiP documents are created by using `Doxygen`.

Man pages

Man pages will be installed at **PIP_INSTALL_DIR**/share/man.

```
$ man -M PIP_INSTALL_DIR/share/man 7 libpip
```

Or, use the pip-man command (from v2).

```
$ PIP_INSTALL_DIR/bin/pip-man 7 libpip
```

The above two exammples will show you the same document you are reading.

PDF

`PDF documents` will be installed at **PIP_INSTALL_DIR**/share/pdf.

HTML

`HTML documents` will be installed at **PIP_INSTALL_DIR**/share/html.

Getting Started

Compile and link your PiP programs

- pipcc(1) command (since v2)

You can use pipcc(1) command to compile and link your PiP programs.

```
$ pipcc -Wall -O2 -g -c pip-prog.c
$ pipcc -Wall -O2 -g pip-prog.c -o pip-prog
```

Run your PiP programs

- pip-exec(1) command (piprun(1) in PiP v1)

Let's assume that you have a non-PiP program(s) and wnat to run as PiP tasks. All you have to do is to compile your program by using the above pipcc(1) command and to use the pip-exec(1) command to run your program as PiP tasks.

```
$ pipcc myprog.c -o myprog
$ pip-exec -n 8 ./myprog
$ ./myprog
```

In this case, the pip-exec(1) command becomes the PiP root and your program runs as 8 PiP tasks. Note that the 'myprog.c' may or may not call any PiP functions. Your program can also run as a normal program (not as a PiP task) without using the pip-exec(1) command. In either case, your proghrams must be compiled and linked by using the pipcc(1) command described above.

You may write your own PiP programs whcih includes the PiP root programming. In this case, your program can run without using the pip-exec(1) command.

If you get the following message when you try to run your program;

```
PiP-ERR(19673) './myprog' is not PIE
```

Then this means that the 'myprog' (having PID 19673) is not compiled by using the pipcc(1) command properly. You may check if your program(s) can run as a PiP root and/or PiP task by using the pip-check(1) command (from v2);

```
$ pip-check a.out
a.out : Root&Task
```

Above example shows that the 'a.out' program can run as a PiP root and PiP tasks.

- pips(1) command (from v2)

  Similar to the Linux `ps` command, you can see how your PiP program(s) is (are) running by using the `pips(1)` command. `pips` can accept 'a', 'u' and 'x' options just like the `ps` command.

  $ pips [a][u][x] [PIPS-OPTIONS] [-] [PATTERN ..]

  List the PiP tasks via the 'ps' command;

  $ pips –ps [PATTERN ..]

  or, show the activities of PiP tasks via the 'top' command;

  $ pips –top [PATTERN ..]

  Additionally you can kill all of your PiP tasks by using the same pips(1) command;

  $ pips -s KILL [PATTERN ..]

Debugging your PiP programs by the pip-gdb command

The following procedure attaches all PiP tasks and PiP root which created those tasks. Each PiP task is treated as a GDB inferior in PiP-gdb. Note that PiP-glibc and PiP-gdb packages are required to do this.

```
$ pip-gdb
(pip-gdb) attach PID
```

The attached inferiors can be seen by the following GDB command:

```
(pip-gdb) info inferiors
  Num  Description              Executable
  4    process 6453 (pip 2)     /somewhere/pip-task-2
  3    process 6452 (pip 1)     /somewhere/pip-task-1
  2    process 6451 (pip 0)     /somewhere/pip-task-0
* 1    process 6450 (pip root)  /somewhere/pip-root
```

You can select and debug an inferior by the following GDB command:

```
(pip-gdb) inferior 2
[Switching to inferior 2 [process 6451 (pip 0)] (/somewhere/pip-task-0)]
```

When an already-attached program calls 'pip_spawn()' and becomes a PiP root task, the newly created PiP child tasks aren't attached automatically, but you can add empty inferiors and then attach the PiP child tasks to the inferiors. e.g.

```
.... type Control-Z to stop the root task.
^Z
Program received signal SIGTSTP, Stopped (user).

(pip-gdb) add-inferior
Added inferior 2
(pip-gdb) inferior 2
(pip-gdb) attach 1902

(pip-gdb) add-inferior
Added inferior 3
(pip-gdb) inferior 3
(pip-gdb) attach 1903

(pip-gdb) add-inferior
Added inferior 4
(pip-gdb) inferior 4
(pip-gdb) attach 1904

(pip-gdb) info inferiors
  Num  Description               Executable
* 4    process 1904 (pip 2)      /somewhere/pip-task-2
  3    process 1903 (pip 1)      /somewhere/pip-task-1
  2    process 1902 (pip 0)      /somewhere/pip-task-0
  1    process 1897 (pip root)   /somewhere/pip-root
```

You can attach all relevant PiP tasks by:

```
$ pip-gdb -p PID-of-your-PiP-program
```

(from v2)

If the **PIP_GDB_PATH** environment is set to the path pointing to PiP-gdb executable file, then PiP-gdb is automatically attached when an excetion signal (SIGSEGV and SIGHUP by default) is delivered. The exception signals can also be defined by setting the **PIP_GDB_SIGNALS** environment. Signal names (case insensitive) can be concatenated by the '+' or '-' symbol. 'all' is reserved to specify most of the signals. For example, 'ALL-TERM' means all signals excepting SIGTERM, another example, 'PIPE+INT' means SIGPIPE and SIGINT. If one of the specified or default signals is delivered, then PiP-gdb will be attached automatically. The PiP-gdb will show backtrace by default. If users specify **PIP_GDB_COMMAND**, a filename containing some GDB commands, then those GDB commands will be executed by PiP-gdb in batch mode. If the **PIP_STOP_ON_START** environment is set, then the PiP library delivers SIGSTOP to a spawned PiP task which is about to start user program. If its value is a number in decimal, then the PiP task whose PiP-ID is the same with the specified number will be stopped. If the number is minus, then all PiP tasks will be stopped at the very beginning. Do not forget to compile your programs with a debug option.

Mailing List

pip@ml.riken.jp

Publications

Research papers

Atsushi Hori, Min Si, Balazs Gerofi, Masamichi Takagi, Jay Dayal, Pavan Balaji, and Yutaka Ishikawa. "Process-in-process: techniques for practical address-space sharing," In Proceedings of the 27th International Symposium on High-Performance Parallel and Distributed Computing (HPDC '18). ACM, New York, NY, USA, 131-143. DOI: https://doi.org/10.1145/3208040.3208045

Atsushi Hori, Balazs Gerofi, and Yuataka Ishikawa. "An Implementation of User-Level Processes using Address Space Sharing," 2020 IEEE International Parallel and Distributed Processing Symposium Workshops (IPDPSW), New Orleans, LA, USA, 2020, pp. 976-984, DOI: https://doi.org/10.1109/IPDPSW50202.2020.-00161.

Kaiming Ouyang, Min Si, Atsushi Hori, Zizhong Chen and Pavan Balaji. "CAB-MPI: Exploring Interprocess Work Stealing toward Balanced MPI Communication," in SC'20

Commands

- pipcc
- pip-check
- pip-exec
- pip-man
- pip-mode
- pips
- printpipmode

Functions

- pip_abort
- pip_barrier_fin
- pip_barrier_init
- pip_barrier_wait
- pip_exit
- pip_export
- pip_fin
- pip_get_aux
- pip_get_mode
- pip_get_mode_str
- pip_get_ntasks
- pip_get_pipid
- pip_get_system_id
- pip_import
- pip_init
- pip_isa_root
- pip_isa_task
- pip_is_initialized
- pip_is_shared_fd
- pip_is_threaded
- pip_kill
- pip_kill_all_tasks

- pip_named_export
- pip_named_import
- pip_named_tryimport
- pip_set_aux
- pip_sigmask
- pip_signal_wait
- pip_spawn
- pip_spawn_from_func
- pip_spawn_from_main
- pip_spawn_hook
- pip_task_spawn
- pip_trywait
- pip_trywait_any
- pip_wait
- pip_wait_any
- pip_yield

Author

Atsushi Hori

Riken Center for Commputational Science (R-CCS)

Japan

Chapter 2

PiP Commands

2.1 pipcc

C compiler driver for PiP

Synopsis

pipcc [PIP-OPTIONS] [CC-COMMAND-OPTIONS_AND_ARGS]

Parameters

-piproot	the compile (and link) as a PiP root
-piptask	the compile (and link) as a PiP task
-nopip	No PiP related settings will be applied

Note

The **-piproot** and **-piptask** options can be specified at the same time. In this case, the compiled object can be both of PiP root and PiP task. This is also the default behavior when none of them is not specified.

Environment

if `CC` environment is set then `$`(CC) will be used as a C compiler

See Also

pip-exec(1), pip-mode(1)

2.2 pip-check

PiP binary checking program if a progarm can run sa a PiP root and/or PiP task

Synopsis

pip-check [OPTION] PIP-PROG [...]

Parameters

-r	check if a.out can be PiP root
-t	check if a.out can be PiP task
-b	check if a.out can be PiP root and/or PiP task (default)
-v	show reason
-h	show this message

See Also

pipcc

2.3 pip-exec

run program(s) as PiP tasks

Synopsis

pip-exec [OPTIONS] <program> ... [: ...]

Description

Run a program as PiP task(s). Mutiple programs can be specified by separating them with ':' to share the same virtual address space with the `pip-exec` command.

Parameters

-n N	number of tasks
-f FUNC	function name to start
-c CORE	specify the CPU core number to bind core(s)
-r	core binding in the round-robin fashion

See Also

pipcc(1)

2.4 pip-man

show PiP man page

Synopsis

pip-man [MAN-OPT] MAN-TOPIC

Description

Show PiP man pages. It can also accept the man command options.

See Also

man(1)

2.5 pip-mode

Set PiP execution mode

Synopsis

pip-mode [OPTION] [PIP-COMMAND]

Description

The following options are avilable. If no of them specified, then the compiled output file can be used as both PiP root and PiP task.

Parameters

-P	'process' mode
-L	'process:preload' mode
-C	'process:clone' mode
-G	'process:got' mode
-T	'thread' mode
-u	Show usage

See Also

pip-exec
printpipmode

2.6 pips

List or kill running PiP tasks

Synopsis

pips [a][u][x] [PIP-OPTIONS] [-] [PATTERN ..]

Parameters

a|*u*|*x*	similar to the aux options of the Linux ps command
--root	List PiP root(s)
--task	List PiP task(s)
--family	List PiP root(s) and PiP task(s) in family order
--kill	Send SIGTERM to PiP root(s) and task(s)
--signal	Send a signal to PiP root(s) and task(s). This option must be followed by a signal number of name.
--ps	Run the ps Linux command. This option may have `ps` command option(s) separated by comma (,)
--top	Run the top Linux command. This option may have `top` command option(s) separated by comma (,)
-	Simply ignored. This option can be used to avoid the ambiguity of the options.

Description

`pips` is a filter to target only PiP tasks (including PiP root) to show status like the way what the `ps` commands does and send signals to the selected PiP tasks.

Just like the `ps` command, `pips` can take the most familiar `ps` options `a`, `u`, `x`. Here is an example;

```
$ pips
PID   TID   TT       TIME     PIP COMMAND
18741 18741 pts/0    00:00:00 RT  pip-exec
18742 18742 pts/0    00:00:00 RG  pip-exec
18743 18743 pts/0    00:00:00 RL  pip-exec
18741 18744 pts/0    00:00:00 0T  a
```

```
18745 18745 pts/0    00:00:00 0G  b
18746 18746 pts/0    00:00:00 0L  c
18747 18747 pts/0    00:00:00 1L  c
18741 18748 pts/0    00:00:00 1T  a
18749 18749 pts/0    00:00:00 1G  b
18741 18750 pts/0    00:00:00 2T  a
18751 18751 pts/0    00:00:00 2G  b
18741 18752 pts/0    00:00:00 3T  a
```

here, there are 3 `pip-exec` root processes running. Four pip tasks running program 'a' with the ptherad mode, three PiP tasks running program 'b' with the process:got mode, and two PiP tasks running program 'c' with the process:preload mode.

Unlike the `ps` command, two columns 'TID' and 'PIP' are added. The 'TID' field is to identify PiP tasks in pthread execution mode. three PiP tasks running in the pthread mode. As for the 'PiP' field, if the first letter is 'R' then that pip task is running as a PiP root. If this letter is a number from '0' to '9' then this is a PiP task (not root). The number is the least-significant digit of the PiP ID of that PiP task. The second letter represents the PiP execution mode which is common with PiP root and task. 'L' is 'process:preload,' 'C' is 'process:pipclone,', 'G' is 'process:got,' and 'T' is 'thread.'

The last 'COMMAND' column of the `pips` output may be different from what the `ps` command shows, although it looks the same. It represents the command, not the command line consisting of a command and its argument(s). More precisely speaking, it is the first 14 letters of the command. This comes from the PiP's specificity. PiP tasks are not created by using the normal `exec` systemcall and the Linux assumes the same command line with the pip root process which creates the pip tasks.

If users want to have the other `ps` command options other than 'aux', then refer to the `--ps` option described below. But in this case, the command lines of PiP tasks (excepting PiP roots) are not correct.

- `--root` (`-r`) Only the PiP root tasks will be shown.

```
$ pips --root
PID   TID   TT       TIME     PIP COMMAND
18741 18741 pts/0    00:00:00 RT  pip-exec
18742 18742 pts/0    00:00:00 RG  pip-exec
18743 18743 pts/0    00:00:00 RL  pip-exec
```

- `--task` (`-t`) Only the PiP tasks (excluding root) will be shown. If both of `--root` and `--task` are specified, then firstly PiP roots are shown and then PiP tasks will be shown.

```
$ pips --tasks
PID   TID   TT       TIME     PIP COMMAND
18741 18744 pts/0    00:00:00 0T  a
18745 18745 pts/0    00:00:00 0G  b
18746 18746 pts/0    00:00:00 0L  c
18747 18747 pts/0    00:00:00 1L  c
18741 18748 pts/0    00:00:00 1T  a
18749 18749 pts/0    00:00:00 1G  b
18741 18750 pts/0    00:00:00 2T  a
18751 18751 pts/0    00:00:00 2G  b
18741 18752 pts/0    00:00:00 3T  a
```

- `--family` (`-f`) All PiP roots and tasks of the selected PiP tasks by the `PATTERN` optional argument of `pips`.

```
$ pips - a
PID   TID   TT       TIME     PIP COMMAND
18741 18744 pts/0    00:00:00 0T  a
18741 18748 pts/0    00:00:00 1T  a
18741 18750 pts/0    00:00:00 2T  a
$ pips --family a
PID   TID   TT       TIME     PIP COMMAND
18741 18741 pts/0    00:00:00 RT  pip-exec
18741 18744 pts/0    00:00:00 0T  a
18741 18748 pts/0    00:00:00 1T  a
18741 18750 pts/0    00:00:00 2T  a
```

In this example, "pips - a" (the - is needed not to confused the command name `a` as the `pips` option) shows the PiP tasks which is derived from the program `a`. The second run, "pips --family a," shows the PiP tasks of `a` and their PiP root (`pip-exec`, in this example).

- `--kill` (`-k`) Send SIGTERM signal to the selected PiP tasks.
- `--signal` (`-s`) `SIGNAL` Send the specified signal to the selected PiP tasks.
- `--ps` (`-P`) This option may be followed by the `ps` command options. When this option is specified, the PIDs of selected PiP tasks are passed to the `ps` command with the specified `ps` command options, if given.
- `--top` (`-T`) This option may be followed by the `top` command options. When this option is specified, the PIDs of selected PiP tasks are passed to the `top` command with the specified `top` command options, if given.
- `PATTERN` The last argument is the pattern(s) to select which PiP tasks to be selected and shown. This pattern can be a command name (only the first 14 characters are effective), PID, TID, or a Unix (Linux) filename matching pattern (if the fnmatch Python module is available).

```
$ pips - *-*
PID   TID   TT       TIME     PIP COMMAND
18741 18741 pts/0    00:00:00 RT  pip-exec
18742 18742 pts/0    00:00:00 RG  pip-exec
18743 18743 pts/0    00:00:00 RL  pip-exec
```

Note

`pips` collects PiP tasks' status by using the Linux's `ps` command. When the `--ps` or `--top` option is specified, the `ps` or `top` command is invoked after invoking the `ps` command for information gathering. This, however, may result some PiP tasks may not appear in the invoked `ps` or `top` command when one or more PiP tasks finished after the first `ps` command invocation. The same situation may also happen with the `--kill` or --signal option.

2.7 printpipmode

Print current PiP mode

Synopsis

printpipmode

See Also

pip-mode

Chapter 3

PiP Functions