

User Guide to `lunix`, Comprehensive Unix API Module for Lua

William Ahern

May 11, 2017

Contents

Contents	i
1 About	1
2 Dependencies	2
2.1 Operating Systems	2
2.2 Libraries	2
2.2.1 Lua 5.1, 5.2, 5.3	2
2.3 GNU Make	2
3 Installation	3
3.1 Building	3
3.1.1 Targets	3
3.2 Installing	3
3.2.1 Targets	4
4 Usage	5
4.1 Modules	5
4.1.1 <code>unix</code>	5
<code>environ[]</code>	5
<code>accept</code>	5
<code>arc4random</code>	5
<code>arc4random_buf</code>	6
<code>arc4random_stir</code>	6
<code>arc4random_uniform</code>	6
<code>bind</code>	6
<code>chdir</code>	6
<code>chmod</code>	6
<code>chown</code>	6
<code>chroot</code>	6
<code>clock_gettime</code>	6
<code>closedir</code>	7
<code>connect</code>	7
<code>dup</code>	7
<code>dup2</code>	7
<code>dup3</code>	7
<code>execve</code>	7
<code>execl</code>	7
<code>execvp</code>	8
<code>_exit</code>	8
<code>exit</code>	8
<code>fcntl</code>	8
<code>fdatasync</code>	8
<code>fdopen</code>	8
<code>fdopendir</code>	8
<code>fdup</code>	8
<code>fileno</code>	9
<code>flockfile</code>	9
<code>fstat</code>	9
<code>fsync</code>	9

ftrylockfile	9	mkpath	16
funlockfile	9	open	16
fopen	9	openat	17
fopenat	9	opendir	17
fpipe	9	pipe	17
fork	10	poll	17
gai_strerror	10	posix_fadvise	18
getaddrinfo	10	posix_fallocate	18
getnameinfo	11	posix_openpt	18
getegid	11	posix_fopenpt	18
getenv	11	pread	18
geteuid	11	ptsname	18
getgroups	11	pwrite	18
getmode	11	raise	18
getgid	11	read	18
getgrnam	12	readdir	18
getgroups	12	recv	19
gethostname	12	recvfrom	19
getifaddrs	12	recvfromto	19
getpeername	13	rename	20
getpgid	13	renameat	20
getpgrp	13	rewinddir	20
getpid	13	rmdir	20
getppid	13	S_ISBLK	20
getprogname	13	S_ISCHR	20
getpwnam	14	S_ISDIR	20
getrlimit	14	S_ISFIFO	20
getrusage	14	S_ISREG	21
getsockname	14	S_ISLNK	21
gettimeofday	15	S_ISSOCK	21
getuid	15	send	21
grantpt	15	sendto	21
ioctl	15	sendtofrom	21
isatty	15	setegid	22
issetugid	15	seteuid	22
kill	15	setenv	22
lchown	15	setgid	22
link	15	setgroups	22
lockf	15	setlocale	22
lseek	15	setpgid	22
lstat	16	setrlimit	23
mkdir	16	setsid	23
mkdirat	16	setsockopt	23
mkfifo	16	setuid	23
mkfifoat	16	sigaction	23

sigfillset	23	tcsetpgrp	26
sigemptyset	24	timegm	27
sigaddset	24	truncate	27
sigdelset	24	tzset	27
sigismember	24	umask	27
sigprocmask	24	uname	27
sigtimedwait	24	unlink	28
sigwait	25	unlinkat	28
sleep	25	unlockpt	28
stat	25	unsetenv	28
strerror	26	wait	28
strsignal	26	waitpid	28
symlink	26	write	28
tcgetpgrp	26		
4.1.2 unix.dir	28		
dir:files	28	dir:rewind	29
dir:read	28	dir:close	29
5 Appendix			30
5.1 setrecvaddr	30		
5.2 strevents	31		

1 About

`lunix` is a bindings library module to common Unix system APIs. The module is regularly tested with Linux/glibc, Linux/musl, OS X, FreeBSD, NetBSD, OpenBSD, Solaris, and AIX. The best way to describe it is in contradistinction to `luaposix`, the most popular bindings module for Unix APIs in Lua.

Thread-safety Unlike `luaposix`, it strives to be as thread-safe as possible on the host platform. Interfaces like `strerror_r` and `O_CLOEXEC` are used throughout. The module even includes a novel solution for the inherently non-thread-safe `umask` system call, where calling `umask` from one thread might result in another thread creating a file with unsafe or unexpected permissions.

POSIX Extensions Unlike `luaposix`, the library does not restrict itself to POSIX, and emulates an interface when not available natively on a supported platform. For example, the library provides `arc4random` (absent on Linux and Solaris), `clock_gettime` (absent on OS X), and a thread-safe `timegm` (absent on Solaris).

Leak-safety Unlike `luaposix`, the library prefers dealing with `FILE` handles rather than raw integer descriptors. This helps to mitigate and prevent leaks or double-close bugs—a common source of problems in, e.g., asynchronous applications. Routines like `chdir`, `stat`, and `opendir` transparently accept string paths, `FILE` handles, `DIR` handles, and raw integer descriptors.

2 Dependencies

2.1 Operating Systems

`linux` targets modern POSIX-conformant and POSIX-aspiring systems. But unlike `luaposix` it branches out to implement common GNU and BSD extensions. All interfaces are available on all supported platforms, regardless of whether the platform provides a native interface.

I try to regularly compile and test the module against recent versions of OS X, Linux/glibc, Linux/musl, FreeBSD, NetBSD, OpenBSD, Solaris, and AIX.

2.2 Libraries

2.2.1 Lua 5.1, 5.2, 5.3

`linux` targets Lua 5.1 and above.

2.3 GNU Make

The Makefile requires GNU Make, usually installed as `gmake` on platforms other than Linux or OS X. The actual `Makefile` proxies to `GNUmakefile`. As long as `gmake` is installed on non-GNU systems you can invoke your system's `make`.

3 Installation

The module is composed of a single C source file to simplify compilation across environments. Because there several extant versions of Lua often used in parallel on the same system, there are individual targets to build and install the module for each supported Lua version. The targets `all` and `install` will attempt to build and install both Lua 5.1 and 5.2 modules.

Note that building and installation can be accomplished in a single step by simply invoking one of the install targets with all the necessary variables defined.

3.1 Building

There is no separate `./configure` step required.¹ System introspection and feature detection occurs during compile-time. The “`configure`” make target can be used to cache the build environment so one needn’t continually use a long command-line invocation.

All the common GNU-style compiler variables are supported, including `CC`, `CPPFLAGS`, `CFLAGS`, `LDFLAGS`, and `SOFLAGS`. Note that you can specify the path to Lua 5.1, Lua 5.2, and Lua 5.3 include headers at the same time in `CPPFLAGS`; the build system will work things out to ensure the correct headers are loaded when compiling each version of the module.

3.1.1 Targets

`all`
Build modules for Lua 5.1 and 5.2.

`all5.1`
Build Lua 5.1 module.

`all5.2`
Build Lua 5.2 module.

`all5.3`
Build Lua 5.3 module.

3.2 Installing

All the common GNU-style installation path variables are supported, including `prefix`, `bindir`, `libdir`, `datadir`, `includedir`, and `DESTDIR`. These additional path variables are also allowed:

`lua51path`
Install path for Lua 5.1 modules, e.g. `$(prefix)/share/lua/5.1`

¹Optional autoconf configuration is currently being tested. Run `./bootstrap` to build the `./configure` script

`lua51cpath`

Install path for Lua 5.1 C modules, e.g. `$(prefix)/lib/lua/5.1`

`lua52path`

Install path for Lua 5.2 modules, e.g. `$(prefix)/share/lua/5.2`

`lua52cpath`

Install path for Lua 5.2 C modules, e.g. `$(prefix)/lib/lua/5.2`

`lua53path`

Install path for Lua 5.3 modules, e.g. `$(prefix)/share/lua/5.3`

`lua53cpath`

Install path for Lua 5.3 C modules, e.g. `$(prefix)/lib/lua/5.3`

3.2.1 Targets

`install`

Install modules for Lua 5.1 and 5.2.

`install5.1`

Install Lua 5.1 module.

`install5.2`

Install Lua 5.2 module.

`install5.3`

Install Lua 5.3 module.

4 Usage

4.1 Modules

4.1.1 `unix`

At present `unix` provides a single module of routines.

`environ[]`

Binding to the process-global `environ` array using metamethods.

`__index`

Utilizes the internal `getenv` binding.

`__newindex`

Utilizes the internal `setenv` binding.

`__pairs`

Takes a snapshot of the `environ` table to be used by the returned iterator for key–value loops. *Other than Solaris¹, no system supports thread-safe access of the `environ` global.*

`__ipairs`

Similar to `__pairs`, but the iterator returns an index integer as the key followed by the environment variable as a single string—“FOO=BAR”.

`__call`

Identical to the `__pairs` metamethod, to be used to create an iterator directly as Lua 5.1 doesn’t support `__pairs`.

`accept(file[, flags])`

FIXME.

`arc4random()`

Returns a cryptographically strong uniformly random 32-bit integer as a Lua number. On Linux the `RANDOM.UUID sysctl` feature is used to seed the generator if available; or on more recent Linux and Solaris kernels the `getrandom` interface.² This avoids fiddling with file descriptors, and also works in a chroot jail. On other platforms without a native `arc4random` interface, such as Solaris 11.2 or earlier, the implementation must resort to `/dev/urandom` for seeding.

Note that unlike the original implementation on OpenBSD, `arc4random` on some older platforms (e.g. FreeBSD prior to 10.10) seeds itself from `/dev/urandom`. This could cause problems in chroot jails.

¹See https://blogs.oracle.com/pgdh/entry/caring_for_the_environment_making

²Some Linux distributions, such as Red Hat, disable `sysctl`.

`arc4random_buf(n)`

Returns a string of length *n* containing cryptographically strong random octets using the same CSPRNG underlying `arc4random`.

`arc4random_stir()`

Stir the `arc4random` entropy pool using the best available resources. This normally should be unnecessary.

`arc4random_uniform([n])`

Returns a cryptographically strong uniform random integer in the interval $[0, n - 1]$ where $n \leq 2^{32}$. If *n* is omitted the interval is $[0, 2^{32} - 1]$ and effectively behaves like `arc4random`.

`bind(file[, sockaddr])`

FIXME.

`chdir(dir)`

If *dir* is a string, attempts to change the current working directory using `chdir`. Otherwise, if *dir* is a FILE handle referencing a directory, or an integer file descriptor referencing a directory, attempts to change the current working directory using `fchdir`.

Returns `true` on success, otherwise returns `false`, an error string, and an integer system error.

`chmod(file, mode)`

file may be either be a string path for use with `chmod`, or a FILE handle or integer file descriptor for use with `fchmod`. *mode* may be an integer value or symbolic string.

Returns `true` on success, otherwise returns `false`, an error string, and an integer system error.

`chown(file[, uid][, gid])`

file may be either be a string path for use with `chown`, or a FILE handle or integer file descriptor for use with `fchown`. *uid* and *gid* may be integer values or symbolic string names.

Returns `true` on success, otherwise returns `false`, an error string, and an integer system error.

`chroot(path)`

Attempt to `chroot` to the specified string *path*.

Returns `true` on success, otherwise returns `false`, an error string, and an integer system error.

`clock_gettime(id)`

id should be the string “realtime” or “monotonic”, or the integer constant `CLOCK_REALTIME` or `CLOCK_MONOTONIC`.

Returns a time value as a Lua floating point number, otherwise returns `nil`, an error string, and an integer system error.

`closedir(dir)`

Closes the DIR handle, releasing the underlying file descriptor.

`connect(file[, sockaddr])`

FIXME.

`dup(file[, flags])`

file may be either a FILE handle or integer file descriptor. *flags* is an optional file status flags integer. If available, `F_DUPFD_CLOEXEC` is used to ensure atomic setting of any `O_CLOEXEC` flag.

Returns an integer descriptor on success, otherwise `nil`, an error string, and an integer system error, an error string, and an integer system error.

`dup2(file, file[, flags])`

file may be either a FILE handle or integer file descriptor. *flags* is an optional file status flags integer. If available, either `dup3` or `F_DUP2FD_CLOEXEC` is used to ensure atomic setting of any `O_CLOEXEC` flag.

Returns an integer descriptor on success, otherwise `nil`, an error string, and an integer system error.

`dup3(file, file, flags)`

Like `dup2`, except *flags* is not optional. This binding will not exist if `dup3` was not available at compile-time, whereas the `dup2` binding is best-effort regarding atomically setting `O_CLOEXEC`.

`execve(path[, argv][, env])`

Executes *path*, replacing the existing process image. *path* should be an absolute pathname as the `$PATH` environment variable is not used. *argv* is a table or ipairs-iterable object specifying the argument vector to pass to the new process image. Traditionally the first such argument should be the basename of *path*, but this is not enforced. If absent or empty the new process image will be passed an empty argument vector. *env* is a table or ipairs-iterable object specifying the new environment. If absent or empty the new process image will contain an empty environment.

On success never returns. On failure returns `false`, an error string, and an integer system error.

`exec1(path, ...)`

Executes *path*, replacing the existing process image. The `$PATH` environment variable is not used. Any subsequent arguments are passed to the new process image. The new process image inherits the current environment table.

On success never returns. On failure returns `false`, an error string, and an integer system error.

`execlp(file, ...)`

Executes *file*, replacing the existing process image. The `$PATH` environment variable is used to search for *file*. Any subsequent arguments are passed to the new process image. The new process image inherits the current environment table.

On success never returns. On failure returns **false**, an error string, and an integer system error.

`execvp(file[, argv])`

Executes *file*, replacing the existing process image. The `$PATH` environment variable is used to search for *file*. Any subsequent arguments are passed to the new process image. The new process image inherits the current environment table.

On success never returns. On failure returns **false**, an error string, and an integer system error.

`_exit([status])`

Exits the process immediately without first flushing and closing open streams, or calling `atexit` handlers. If *status* is boolean **true** or **false**, exits with `EXIT_SUCCESS` or `EXIT_FAILURE`, respectively. Otherwise, *status* is an optional integer status value which defaults to 0 (`EXIT_SUCCESS`).

`exit([status])`

Like `_exit`, but first flushes and closes open streams, and calls `atexit` handlers.

`fcntl(file|dir|fd, ...)`

FIXME

`fdatasync(file|dir|fd)`

FIXME

`fdopen(file|dir|fd[, mode])`

FIXME

`fdopendir(file|dir|fd)`

FIXME

`fdup(file[, flags])`

file may be either a FILE handle or integer file descriptor. *flags* is an optional integer or symbolic mode.

Returns a FILE handle on success, otherwise returns **nil**, an error string, and an integer system error.

`fileno(file|dir|fd)`

Resolves the specified FILE handle or DIR handle to an integer file descriptor. An integer descriptor is returned as-is.

`flockfile(fh)`

Locks the FILE handle *fh*, blocking the current thread if already locked. Returns **true**.

This function only works on FILE handles and not DIR handlers or integer descriptors.

`fstat(path|file|dir|fd [, field ...])`

See **stat**.

`fsync(file|dir|fd)`

FIXME

`ftrylockfile(fh)`

Attempts to lock the FILE handle *fh*. Returns **true** on success or **false** if *fh* was locked by another thread.

`funlockfile(fh)`

Unlocks the FILE handle *fh*. Returns **true**.

`fopen(path|file|dir|fd [, mode] [, perm])`

Similar to the **open** binding except returns a FILE handle. Unlike Lua's **io.open** routine, any string *mode* is passed directly to the **fopen** interface and thus supports all local extensions. One consequence of this is that if *mode* is a string then *perm* is ignored. To ensure that **fopen** is invoked both *path* and *mode* should be strings, otherwise the semantics are emulated by translating *mode* to an integer bitfield and invoking **open**.

Returns a FILE handle on success, otherwise **nil**, an error string, and an integer system error.

The difference between calling **fopen** versus **open + fdopen** is that the **fopen** binding ensures that a descriptor is not leaked if **fdopen** fails or throws an exception.

`fopenat(path [, mode] [, perm])`

See **openat** and **fopen**.

Returns a FILE handle on success, otherwise **nil**, an error string, and an integer system error.

`fpipe([mode])`

FIXME.

`fork()`

Forks a new process. On success returns the PID of the new process in the parent and the integer 0 in the child. Otherwise returns **false**, an error string, and an integer system error.

`gai_strerror(error)`

Returns an error string corresponding to the specified EAI integer *error*.

`getaddrinfo(host, [port], [hints][, field ...])`

Returns an iterator over the addresses resolved for *host* and *port*. If a resolution error occurred, returns **nil**, an error string, an EAI error integer, and a system error integer (if EAI error is EAI_SYSTEM).

host should be a string host name. *port* is an optional port number which defaults to 0.

hints is an optional table controlling the manner and scope of resolution. For example, if a family is not specified the resolver will return a unique result for each address-family combination supported by the system. The following fields are supported:

.flags

Bitwise or of one of more of AI_PASSIVE, AI_CANONNAME, AI_NUMERICHOST, AI_NUMERICSERV, AI_ADDRCONFIG, AI_V4MAPPED (if supported), and AI_ALL (if supported).

.family

AF_UNSPEC, AF_INET, or AF_INET6.

.socktype

SOCK_DGRAM, SOCK_STREAM, or SOCK_SEQPACKET (if supported).

.protocol

IPPROTO_IP, IPPROTO_IPV6, IPPROTO_TCP, IPPROTO_UDP, IPPROTO_ICMP, or IPPROTO_RAW (if supported).

If no fields are specified, the iterator returns a table with the following fields:

.family

See description of family for hints table.

.socktype

See description of socktype for hints table.

.protocol

See description of protocol for hints table.

.addr

IP address as human-readable string.

.canonical

Canonical hostname of IP address (if ALCANONNAME flag was specified in hints table).

.port

Integer port.

If fields are specified, the iterator returns a list of fields in the order specified.

`getnameinfo(sockaddr[, flags])`

FIXME.

getegid()

Returns the effective process GID as a Lua number.

getenv(name)

Returns the value of the environment variable *name* as a string, or `nil` if it does not exist.

*Not thread-safe on any system other than Solaris³ and NetBSD⁴. Linux/glibc **getenv** is thread-tolerant as pointers returned from **getenv** will remain valid throughout the lifetime of the process, but Linux/glibc will write over existing values on update so concurrent use with **setenv** could lead to inconsistent views.*

geteuid()

Returns the effective process UID as a Lua number.

getgroups()

Queries supplement group list. On success returns an array of supplement GIDs. Otherwise returns `nil`, an error string, and an integer system error.

getmode(mode[, omode])

The **getmode** interface derives from the routine so-named in almost every `chmod(1)` utility implementation and which exposes the parser for symbolic file permissions.

mode should be a symbolic mode value with a valid syntax as described by POSIX within the `chmod(1)` utility man page. If specified, *omode* should be an integer or a string in decimal, hexadecimal, or octal notation, and represents the original mode value used by the symbolic syntax for inheritance.

getgid()

Returns the real process GID as a Lua number.

³See https://blogs.oracle.com/pgdh/entry/caring_for_the_environment_making

⁴NetBSD provides `getenv_r(3)`

`getgrnam(grp[, ...])`

grp is an integer GID or string symbolic group name suitable for use by either `getgrgid(3)` or `getgrnam(3)`, respectively.

If no other arguments are specified, on success returns a table with the following fields

.name

Symbolic group name as a string, or `nil` if absent.

.passwd

Password information as a string, or `nil` if absent.

.gid

GID as integer.

.mem

Array of supplementary group names, or `nil` if absent.

If additional arguments are given, on success each field specified (as named above) is returned as part of the return value list. “members” may be used as an alternative to “mem”. Note that the return value may be `nil` if the field was absent.

If no group was found, returns `nil` followed by the error string “no such group”.

If a system error occurred, returns `nil`, an error string, and an integer system error.

`getgroups()`

Returns table of supplementary GIDs on success, otherwise `nil`, an error string, and an integer system error.

`gethostname()`

Returns system hostname as string on success, otherwise `nil`, an error string, and an integer system error.

`getifaddrs([...])`

Returns an iterator over the current system network interfaces on success. If a system error occurred, returns `nil`, an error string, and an integer system error.

If no arguments are specified, each invocation of the iterator returns a table with the following fields

.name

Interface symbolic name as a string.

.flags

Interface flags as an integer bit field.

.family

Interface address family as an integer.

.addr

Interface address as a string, or `nil` if of an unknown address family.

.netmask

Interface address netmask as a string, or `nil` if absent or of an unknown address family.

.prefixlen

Interface address prefixlen as an integer, or `nil` if absent or of an unknown address family.

.dstaddr

Interface destination address if point-to-point, or `nil` if absent or of an unknown address family.

.broadcast

Interface broadcast address, or `nil` if absent or of an unknown address family.

If arguments are given, each field specified (as named above) is returned as part of the return value list on every invocation of the iterator.

`getpeername(file)`

FIXME.

`getpgid()`

FIXME.

`getpgrp()`

FIXME.

`getpid()`

Returns the process ID as a Lua number.

`getppid()`

Returns the parent process ID as a Lua number.

`getprogname()`

Returns the program name as a string, otherwise `nil`, an error string, and an integer system error.

`getpwnam(usr[, ...])`

usr is an integer UID or string symbolic user name suitable for use by either `getpwuid(3)` or `getpwnam(3)`, respectively.

If no other arguments are specified, on success returns a table with the following fields

.name

Symbolic user name as a string, or `nil` if absent.

.passwd

Password information as a string, or `nil` if absent.

.uid

UID as integer.

.gid

Primary GID as integer.

.dir

Home directory path, or `nil` if absent.

.shell

Login shell path, or `nil` if absent.

.gecos

Additional user information, or `nil` if absent.

If additional arguments are given, on success each field specified (as named above) is returned as part of the return value list. Note that the return value may be `nil` if the value was empty in the database.

If no user was found, returns `nil` followed by the error string “no such user”.

If a system error occurred, returns `nil`, an error string, and an integer system error.

`getrlimit([what])`

FIXME.

`getrusage([who])`

FIXME.

`getsockname(file)`

FIXME.

`gettimeofday([ints])`

Returns the current time as a Lua floating point number or, if *ints* is **true**, as two integers representing seconds and microseconds.

On failure returns **nil**, an error string, and an integer system error.

`getuid()`

Returns the real process UID as a Lua number.

`grantpt(file)`

FIXME.

`ioctl(file, ...)`

FIXME.

`isatty(file)`

FIXME.

`issetugid()`

Returns **true** if the process environment is considered unsafe because of `setuid`, `setgid`, or similar operations, otherwise **false**.

`kill(pid, signo)`

Sends signal *signo* to process or process group *pid*. Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`lchown(path1[, uid][, gid])`

FIXME.

`link(path1, path2)`

Creates a new directory entry at *path2* as a hard link to *path1*.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`lockf(file, cmd[, size])`

FIXME.

`lseek(file, offset, whence)`

FIXME.

`lstat(path[, field ...])`

Identical to `stat`, except only accepts string paths and uses the `lstat` system call.

`mkdir(path[, mode])`

Create a new directory at *path*. *mode*, if specified, should be a symbolic mode string following the POSIX syntax as described by the `chmod(1)` utility man page. Otherwise, *mode* defaults to 0777. In either case, *mode* is masked by the process umask.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`mkdirat(fd|file|dir, path[, mode])`

Similar to `mkdir` except interprets *path* relative to the specified open directory file descriptor, *fd*, or a FILE handle or DIR handle that wraps such a descriptor.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`mkfifo(path[, mode])`

Create a FIFO at *path*. *mode*, if specified, should be an integer bitfield or symbolic string following the POSIX syntax as described by the `chmod(1)` utility man page. Otherwise, *mode* defaults to 0666.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`mkfifoat(fd|file|dir, path[, mode])`

Similar to `mkfifo` except interprets *path* relative to the specified open directory file descriptor, *fd*, or a FILE handle or DIR handle that wraps such a descriptor.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`mkpath(path[, mode][, imode])`

Like `mkdir`, but also creates intermediate directories if missing. *imode* is the mode for intermediate directories. Like *mode* it is restricted by the process umask, but unlike *mode* the user write bit is unconditionally set to ensure the full path can be created.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`open(path|file|dir|fd[, mode][, perm])`

Open the specified file. Normally a string path is specified. If a FILE handle, DIR handle, or file descriptor are specified the library will attempt to use system extensions to create a new file descriptor. Unlike `dup` this descriptor will not share status flags or file position cursors. (This will only work on Linux, NetBSD, and Solaris with `procfs` support; or on macOS using the special “/.vol” namespace. macOS and Solaris only support re-opening file system objects this way, not pipes or sockets. The BSD “/dev/fd” namespace has semantics equivalent to `dup`, except on Linux where “/dev/fd” is a symlink to “/proc/self/fd”.)

mode specifies the open flags as an integer bitfield (e.g. `O_CREAT|O_RDWR, O_RDONLY|O_CLOEXEC`, etc) or a symbolic string (e.g. `"w+x"`). If unspecified defaults to 0, which normally equals `O_RDONLY`.

perm specifies the file creation mode bits as an integer bitfield or a symbolic string (e.g. `"ug+rw,o+r"`, equivalent to `0644`). If unspecified defaults to `0666`.⁵

Returns an integer file descriptor on success, otherwise `nil`, an error string, and an integer system error.

`openat(file|dir|fd, path[, mode][, perm])`

Similar to `open` except the first parameter should resolve to an open directory file descriptor.

Returns an integer file descriptor on success, otherwise `nil`, an error string, and an integer system error.

`opendir(path|file|dir|fd)`

Creates a DIR handle for reading directory entries. Caller may specify a path string, a Lua FILE handle, another DIR handle, or an integer descriptor. In the latter three cases, the underlying descriptor is duplicated using `dup3` (if available) or `dup2` because there's no safe way to steal the descriptor from existing FILE or DIR handles. But it's not a good idea to mix reads between the two original and duplicated descriptors as they will normally share the same open file entry in the kernel, including the same position cursor.⁶

Returns a DIR handle on success, otherwise `nil`, an error string, and an integer system error.

`pipe(mode)`

FIXME.

`poll(fds[, timeout])`

Polls on the file descriptor events specified by *fds*. Following the `luaposix` API, *fds* is a table indexed by integer file descriptors, each table entry another table with the integer bitmak field "events" specifying desired events to poll. *timeout* is an optional floating-pointing value specifying the number of seconds (and fractions thereof) to block. A missing or `nil` timeout will wait forever; a negative value is assumed to be the result of arithmetic underflow and effects a 0-second timeout.⁷

On success updates the "revents" field for each descriptor entry in *fds* and returns the number of ready descriptors; otherwise `nil`, an error string, and an integer system error.

```
1 local fds = { [fd] = { events = unix.POLLIN } }
  local nr = assert(unix.poll(fds, 3.0))
3 for fd, t in pairs(fds) do
  -- see Appendix for strevents implementation
5   print(fd, strevents(t.events), strevents(t.revents))
end
```

⁵Note that the kernel will apply the process-global file mode creation mask to the permission bits. See `umask`.

⁶In the future may add ability to open `/proc/self/fd` or `/dev/fd` entries, which should create a new open file entry.

⁷In other words, do not pass `-1` as-if directly calling the underlying system call.

`posix.fadvise(file, offset, len, advice)`

FIXME.

`posix.fallocate(file, offset, len)`

FIXME.

`posix.openpt([flags])`

FIXME.

`posix.fopenpt([flags])`

FIXME.

`pread(file, size, offset)`

Reads up to *size* bytes of data from *file* at *offset*. *file* may be either a FILE handle or integer file descriptor.

Returns a string on success, otherwise `nil`, an error string, and an integer system error.

`ptsname(file)`

FIXME.

`pwrite(file, data, offset)`

Writes *data* to *file* at *offset*. *file* may be either a FILE handle or integer file descriptor.

Returns an integer representing the number of bytes written (which may be less than `#data`) on success, otherwise `nil`, an error string, and an integer system error.

`raise(signo)`

Sends signal *signo* to calling thread. Returns `true` on success, otherwise `false`, an error string, and an integer system error.

`read(file, size)`

Reads up to *size* bytes of data from *file*. *file* may be either a FILE handle or integer file descriptor.

Returns a string on success, otherwise `nil`, an error string, and an integer system error.

`readdir(dir[, field ...])`

Reads the next directory entry. If no field arguments are specified, on success returns a table with the following fields

.name

Name of file.

.ino

Inode of file.

.type

A numeric value describing the file type, similar to the “mode” field returned by `stat`, except without any permission bits present. You can pass this value to `S_ISREG`, `S_ISDIR`, `S_ISFIFO`, etc.

Available on Linux and BSD derivatives, but, e.g., will be `nil` on Solaris.

If additional arguments are given, on success each field specified (as named above) is returned as part of the return value list. Note that the return value may be `nil` if the value was unavailable.

If the end of directory entries has been reached, returns `nil`.

If a system error occurred, returns `nil`, an error string, and an integer system error.

`recv(file, size[, flags])`

Receives up to *size* bytes of data from the socket *file*. *file* may be either a FILE handle or integer file descriptor. *flags* is an optional integer containing bitwise socket receive flags (e.g. `MSG_WAITALL`).

Returns a string on success, otherwise `nil`, an error string, and an integer system error.

`recvfrom(file, size[, flags])`

Like `recv`. Returns a string and `sockaddr` on success; otherwise `nil`, an error string, and an integer system error.

`recvfromto(file, size[, flags])`

Like `recvfrom`. Returns a string, source `sockaddr`, and destination `sockaddr` on success; otherwise `nil`, an error string, and an integer system error.

This interface is useful for replying to UDP packets with the same source address received on, without having to enumerate, bind, and listen on multiple interfaces. Because interfaces can go down and come up dynamically, binding on multiple interfaces is complex and difficult. See also the reciprocal interface, [sendtofrom](#).

NOTES:

- Not all platforms support this interface. AIX (confirmed 7.1) supports IPv6 but not IPv4.
- The descriptor **must** be initialized by enabling a specialized socket option. The option varies by platform and socket protocol family. See [setrecvaddr](#) example in Appendix.

`rename(from, to)`

Renames the file *from* to *to*. The paths must reside on the same device.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`renameat(fromfd, from, tofd, to)`

Similar to **rename** except interprets the string paths *from* and *to* relative to the respective open directory file descriptor, *fromfd* and *tofd*, respectively. *fromfd* and *tofd* should be integer file descriptors to an open directory, or FILE handles or DIR handles that wrap such descriptors.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`rewinddir(dir)`

Rewinds the DIR handle so the directory entries may be read again.

`rmdir(path)`

Remove the directory at *path*.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`S_ISBLK(mode)`

Tests whether the specified *mode* value—as returned by, e.g., **stat** or **readdir**—represents a block device.

Returns **true** or **false**.

`S_ISCHR(mode)`

Tests whether the specified *mode* value—as returned by, e.g., **stat** or **readdir**—represents a character device.

Returns **true** or **false**.

`S_ISDIR(mode)`

Tests whether the specified *mode* value—as returned by, e.g., **stat** or **readdir**—represents a directory.

Returns **true** or **false**.

`S_ISFIFO(mode)`

Tests whether the specified *mode* value—as returned by, e.g., **stat** or **readdir**—represents a FIFO or pipe.

Returns **true** or **false**.

S_ISREG(*mode*)

Tests whether the specified *mode* value—as returned by, e.g., **stat** or **readdir**—represents a regular file.

Returns **true** or **false**.

S_ISLNK(*mode*)

Tests whether the specified *mode* value—as returned by, e.g., **stat** or **readdir**—represents a symbolic link.

Returns **true** or **false**.

S_ISSOCK(*mode*)

Tests whether the specified *mode* value—as returned by, e.g., **stat** or **readdir**—represents a socket.

Returns **true** or **false**.

send(*file*, *data*[], *flags*)

Sends *data* to the peer on the socket, *file*. *file* may be either a FILE handle or integer file descriptor. *flags* is an optional integer containing bitwise socket send flags (e.g. **MSG_NOSIGNAL**).

Returns an integer representing the number of bytes sent (which may be less than **#data**) on success, otherwise **nil**, an error string, and an integer system error.

sendto(*file*, *data*, *flags*, *to_addr*)

Like **send**. *to_addr* is a **sockaddr** destination address or table convertible to a **sockaddr**

sendtofrom(*file*, *data*, *flags*, *to_addr*, *from_addr*)

Like **send**. *to_addr* is a **sockaddr** destination address or table convertible to a **sockaddr**, and *from_addr* a **sockaddr** source address or table convertible to a **sockaddr**. See also the reciprocal interface, [recvfromto](#).

NOTES:

- Not all platforms support this interface. AIX (confirmed 7.1), NetBSD (confirmed 7.0), and OpenBSD \geq 6.0 (confirmed 6.0) support IPv6 but not IPv4.
- FreeBSD (confirmed 10.1) requires the socket to be bound to the wildcard address, “0.0.0.0” or “::”.
- macOS 10.10 and several prior releases contain a kernel bug which causes a kernel panic if the socket address is not bound as required by FreeBSD. (Oddly macOS copied Linux’s **IP_PKTINFO** interface, not FreeBSD’s **IP_SENDSRCADDR**; but Linux doesn’t have the requirement to bind).

`setegid(gid)`

Set the effective process GID to *gid*. *gid* must be an integer or symbolic group name.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`seteuid(uid)`

Set the effective process UID to *uid*. *uid* must be an integer or symbolic user name.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`setenv(name, value[, overwrite])`

Sets the environment variable *name* to *value*. If the variable already exists then it is not changed unless *overwrite* is **true**. *overwrite* defaults to *true*.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

This function is thread-safe on Solaris and NetBSD. For Linux/glibc see note at `getenv`. FreeBSD, OpenBSD, and Linux/musl are confirmed to be not thread-safe. The status of AIX and OS X is unknown. In general `setenv` should be avoided in multi-threaded environments.

`setgid(gid)`

Set the real process GID to *gid*. *gid* must be an integer or symbolic group name.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`setgroups{ ... }`

Sets the supplement group list. Takes an array of GIDs. On success returns **true**. Otherwise returns **false**, an error string, and an integer system error.

As an extension, group names may be specified instead of integer GIDs. However, an unresolvable group name currently causes an error to be thrown rather than returned. Until this is fixed, use `getgrnam` to explicitly resolve names to GIDs.

`setlocale(category [, locale])`

Set or query the program locale. *category* is an integer constant which specifies the category of localization, and should be one of `LC_ALL`, `LC_COLLATE`, `LC_CTYPE`, `LC_MONETARY`, `LC_NUMERIC`, or `LC_TIME`.

locale can be either a string identifier for the locale, or `nil` to query the specified category. An empty locale string selects the system's native locale.

Returns a locale string identifier on success, otherwise `nil` if the specified *category* and *locale* could not be honored.

`setpgid(pid, pgid)`

FIXME.

`setrlimit(what [, soft] [, hard])`

FIXME.

`setsid()`

Create a new session and process group.

Returns the new process group ID on success, otherwise `nil`, an error string, and an integer system error.

`setsockopt(file, level, type, ...)`

FIXME.

`setuid(uid)`

Set the real process UID to *uid*. *uid* must be an integer or symbolic user name.

Returns `true` on success, otherwise `false`, an error string, and an integer system error.

`sigaction(signo, [action], [oaction])`

Sets or queries the signal disposition for the signal *signo*.

If specified, *action* is a table used to update the signal disposition.

.handler

Currently handler may only be `SIG_DFL`, `SIG_ERR`, or `SIG_IGN`. Lua functions are not currently supported, but may in the future. In the meantime, applications can use `sigtimedwait` to atomically dequeue signals in a thread-safe manner.

.mask

A `sigset_t` userdata object, or the string "*" (see `sigfillset`).

.flags

Bitwise or of one or more of `SA_NOCLDSTOP`, `SA_ONSTACK`, `SA_RESETHAND`, `SA_RESTART`, `SA_SIGINFO`, `SA_NOCLDWAIT`, and `SA_NODEFER`.

Returns `true` on success if *oaction* is `nil` or `false`.

Returns a table on success if *oaction* is `true`. The table describes the signal disposition at the time `sigaction` was initially called.

Otherwise returns `nil`, an error string, and an integer system error.

`sigfillset([set])`

Returns a `sigset_t` userdata object with all bits filled. If *set* is specified should be an existing `sigset_t` userdata object to reuse.

`sigemptyset([set])`

Returns a `sigset_t` userdata object with all bits cleared. If *set* is specified should be an existing `sigset_t` userdata object to reuse.

`sigaddset(set[, signo ...])`

Returns a `sigset_t` userdata object with the specified signals set. If *set* is not a `sigset_t` object, a new, empty `sigset_t` is instantiated and initialized according to whether *set* is `nil`, an integer signal number, an array of integer signal numbers, or the string “*” (filled) or “” (empty). If specified, *signo* and additional arguments should be integer signal numbers to be added to the `sigset_t` object.

`sigdelset(set[, signo ...])`

Like `sigaddset`, but *signo* and subsequent integer signal numbers are cleared from the `sigset_t` object.

`sigismember(set, signo)`

Returns `true` if *signo* is a member of `sigset_t set`, otherwise `false`.

`sigprocmask([how, set[, oset]])`

If *how* and *set* are defined, sets the signal mask of the current process or thread. *how* should be one of `SIG_BLOCK`, `SIG_UNBLOCK`, or `SIG_SETMASK`. *set* should be a `sigset_t` userdata object, or a number, string, or array suitable for initializing a `sigset_t` object as discussed in `sigaddset`.

Returns the old mask as a `sigset_t` userdata object on success, otherwise `nil`, an error string, and an integer system error. *oset* is an optional `sigset_t` userdata object to be reused as the return value, and is first cleared before passing to the system call.

Whether the process or thread mask is set is implementation defined, and varies across platforms. Threaded applications should use `pthread_sigmask`, which is guaranteed to set the mask of the current thread.⁸ Unfortunately, there is no interface which is guaranteed to only set the process mask. New threads inherit the mask of the creating thread, so standard practice is typically to block everything in the main thread while creating new threads.

`sigtimedwait(set[, timeout])`

Atomically clears any pending signal specified in *set* from the pending set of the process *and* thread. If none are pending, waits for *timeout* seconds, or indefinitely if *timeout* is not specified. Fractional seconds are supported.

On success returns an integer signal number cleared from the pending set and an array representing the members of the `siginfo_t` structure (without the “si_” prefix).⁹ On error returns `nil`, an error

⁸Use of `pthread_sigmask` requires linking with `-lpthread` on some platforms and for this reason is presently not supported by `lunix`.

⁹Currently only the `.si_signo` member is copied from `siginfo_t`.

string, and an integer system error. If *timeout* is specified and no signal was cleared before the timeout, the system error will be ETIMEDOUT.

OS X and OpenBSD lack a native `sigtimedwait` implementation. On OS X `linux` uses `sigpending` and `sigwait` to emulate the behavior. However, in a multi-threaded application if another thread clears a signal between `sigpending` and `sigwait` then `sigwait` could block indefinitely. There's no way to solve this race condition.¹⁰ On OpenBSD `sigwait` is only available through `libpthread`, but on OpenBSD `libpthread` must be loaded at process load-time and cannot be brought in as a `dlopen` run-time dependency. Therefore an alternative emulation is used which clears the pending signal by installing a noop signal handler. This is not thread-safe if another thread is also installing a signal handler simultaneously. Threaded applications on these platforms should be mindful of these limitations. The `cqueues` project supports thread-safe signal listening with `kqueue` on both OpenBSD and Mac OS X.

`sigwait(set)`

FIXME.

`sleep(n)`

FIXME.

`stat(path|file|dir|fd[, field ...])`

Stats the specified file. Caller may specify a path string, a Lua FILE handle, a DIR handle (see `opendir`), or an integer descriptor.

If no field arguments are specified, on success returns a table with the following fields

.dev

Device identifier as integer of device containing file.

.ino

Inode identifier as integer.

.mode

Mode—type, permissions, etc—as integer.

.nlink

Link count as integer.

.uid

Owner UID as integer.

.gid

Owner GID as integer.

¹⁰One possible solution is to explicitly `raise` the signal before calling `sigpending`, but this solutions relies on untested assumptions about signal handling on these platforms.

.rdev

Device identifier as integer if character or block special file.

.size

File size as integer.

.atime

Last data access timestamp as floating-point number with sub-second fractional component¹¹.

.mtime

Last data modification timestamp as floating-point number with sub-second fractional component.

.ctime

Last file status change timestamp as floating-point number with sub-second fractional component.

.blksize

File-system-specified preferred I/O block size as integer.

.blocks

Number of blocks allocated for this object as integer.

If field arguments are given, on success each field specified (as named above) is returned as part of the return value list. Note that the return value may be `nil` if the value was unavailable.

On error returns `nil`, an error string, and an integer system error.

strerror(*error*)

Returns an error string corresponding to the specified system *error* integer.

strsignal(*signo*)

Returns a string describing the specified signal number.

symlink(*path1*, *path2*)

Creates a new directory entry at *path2* as a symbolic link to *path1*.

Returns `true` on success, otherwise `false`, an error string, and an integer system error.

tcgetpgrp(*file*)

FIXME.

tcsetpgrp(*file*, *pgid*)

FIXME.

¹¹All platforms currently support timestamps with sub-second precision. However, the underlying filesystem may not record a timestamp with sub-second precision.

`timegm(tm)`

tm is a table of the form returned by the Lua routine `os.date("*t")`. This allows converting a datetime in GMT directly to a POSIX timestamp without having to change the process timezone, which is inherently non-thread-safe.

Returns a POSIX timestamp as a Lua number.

`truncate(file[, size])`

Truncate *file* to *size* bytes (defaults to 0). *file* should be a string path, or **FILE** handle or integer file descriptor.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`tzset()`

Initializes datetime conversion information according to the TZ environment variable, if available.

Return **true**.

`umask([cmask])`

If *cmask* is specified, sets the process file creation mask and returns the previous mask as a Lua number.

If *cmask* is not specified, queries the process umask in a thread-safe manner and returns the mask as a Lua number.

`uname([...])`

If no arguments are given, on success returns a table with the following fields

.sysname

Name of the current system as a string.

.nodename

Name of this node within an implementation-defined communications network as a string.

.release

Release name of the operating system as a string.

.version

Version of the operating system as a string.

.machine

Hardware description of the system as a string.

If additional arguments are given, on success each field specified (as named above) is returned as part of the return value list.

On failure returns **nil**, an error string, and an integer system error.

`unlink(path)`

Deletes the file entry at *path*.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`unlinkat(file|dir|fd, path[, flags])`

Deletes the file entry at *path* relative to the specified open directory file descriptor, *fd*, or a FILE handle or DIR handle that wraps such a descriptor.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

`unlockpt(file)`

FIXME.

`unsetenv(name)`

Deletes the environment variable *name* from the environment table.

Returns **true** on success, otherwise **false**, an error string, and an integer system error.

*This function is thread-safe on Solaris, NetBSD, and Linux/glibc. But see note at **getenv**. Also see note at **setenv**. In general **unsetenv** should be avoided in multi-threaded environments.*

`wait([pid] [, options])`

FIXME.

`waitpid([pid] [, options])`

FIXME.

`write(file, data)`

Writes *data* to *file*. *file* may be either a FILE handle or integer file descriptor.

Returns an integer representing the number of bytes written (which may be less than **#data**) on success, otherwise **nil**, an error string, and an integer system error.

4.1.2 `unix.dir`

The `unix.dir` module implements the prototype for DIR handles, as returned by `unix.opendir`.

`dir:files([field ...])`

Returns an iterator over `unix.readdir(...)`.

`dir:read([field ...])`

Identical to `unix.readdir`.

`dir:rewind()`

Identical to `unix.rewinddir`.

`dir:close()`

Identical to `unix.closedir`.

5 Appendix

5.1 setrecvaddr: Prepare socket for recvfromto

A routine to set the socket options necessary for `recvfromto`. Options should be set before binding the socket to an address as on some platforms (e.g. FreeBSD) any packets received beforehand will not be tagged with the destination address.

```
local unix = require"unix"
2
local function setrecvaddr(fd, family)
4   local type, level

6   if family == unix.AF_INET6 then
       level = unix.IPPROTO_IPV6
8       type = unix.IPV6_RECVPKTINFO or unix.IPV6_PKTINFO
   elseif family == unix.AF_INET
10      level = unix.IPPROTO_IP
       type = unix.IP_RECVSTADDR or unix.IP_PKTINFO
12   end

14   if level and type then
       return unix.setsockopt(fd, level, type, true)
16   else
       local errno = unix.EAFNOSUPPORT
18       return false, unix.strerror(errno), errno
   end
20 end
```

5.2 strevents: Stringify poll events

A routine to translate an integer bitfield of poll events to a string of constant names.

```
local unix = require"unix"
2
local function strconst(c, patt)
4   for k,v in pairs(unix) do
       if c == v and type(k) == "string" and k:match(patt) then
6       return k
       end
8   end
   end
10
local function strevent(event)
12   return strconst(event, "^POLL")
   end
14
local function strevents(events)
16   local t = {}
       for i=0,30 do
18       if events % 2 == 1 then
           t[#t + 1] = strevent(2^i)
20       events = events - 1
           end
22       events = events / 2
       end
24   return #t > 0 and table.concat(t, ", ") or nil
   end
```