

**NAME**

perlpio - perl's IO abstraction interface.

**SYNOPSIS**

```
#define PERLIO_NOT_STDIO 0 /* For co-existence with stdio only */
#include <perlio.h> /* Usually via #include <perl.h> */
```

```
PerlIO *PerlIO_stdin(void);
PerlIO *PerlIO_stdout(void);
PerlIO *PerlIO_stderr(void);
```

```
PerlIO *PerlIO_open(const char *path, const char *mode);
PerlIO *PerlIO_fdopen(int fd, const char *mode);
PerlIO *PerlIO_reopen(const char *path, /* deprecated */
    const char *mode, PerlIO *old);
int PerlIO_close(PerlIO *f);
```

```
int PerlIO_stdoutf(const char *fmt,...)
int PerlIO_puts(PerlIO *f, const char *string);
int PerlIO_putc(PerlIO *f, int ch);
SSize_t PerlIO_write(PerlIO *f, const void *buf, size_t numbytes);
int PerlIO_printf(PerlIO *f, const char *fmt,...);
int PerlIO_vprintf(PerlIO *f, const char *fmt, va_list args);
int PerlIO_flush(PerlIO *f);
```

```
int PerlIO_fill(PerlIO *f);
int PerlIO_eof(PerlIO *f);
int PerlIO_error(PerlIO *f);
void PerlIO_clearerr(PerlIO *f);
```

```
int PerlIO_getc(PerlIO *d);
int PerlIO_ungetc(PerlIO *f, int ch);
SSize_t PerlIO_read(PerlIO *f, void *buf, size_t numbytes);
Size_t PerlIO_unread(PerlIO *f, const void *vbuf, size_t count)
```

```
int PerlIO_fileno(PerlIO *f);
```

```
void PerlIO_setlinebuf(PerlIO *f);
```

```
Off_t PerlIO_tell(PerlIO *f);
```

```

int  PerlIO_seek(PerlIO *f, Off_t offset, int whence);
void  PerlIO_rewind(PerlIO *f);

int  PerlIO_getpos(PerlIO *f, SV *save); /* prototype changed */
int  PerlIO_setpos(PerlIO *f, SV *saved); /* prototype changed */

int  PerlIO_fast_gets(PerlIO *f);
int  PerlIO_has_cntptr(PerlIO *f);
SSize_t PerlIO_get_cnt(PerlIO *f);
char  *PerlIO_get_ptr(PerlIO *f);
void  PerlIO_set_ptrcnt(PerlIO *f, char *ptr, SSize_t count);

int  PerlIO_canset_cnt(PerlIO *f);      /* deprecated */
void  PerlIO_set_cnt(PerlIO *f, int count); /* deprecated */

int  PerlIO_has_base(PerlIO *f);
char  *PerlIO_get_base(PerlIO *f);
SSize_t PerlIO_get_bufsiz(PerlIO *f);

PerlIO *PerlIO_importFILE(FILE *stdio, const char *mode);
FILE  *PerlIO_exportFILE(PerlIO *f, const char *mode);
FILE  *PerlIO_findFILE(PerlIO *f);
void  PerlIO_releaseFILE(PerlIO *f, FILE *stdio);

int  PerlIO_apply_layers(pTHX_ PerlIO *f, const char *mode,
                        const char *layers);
int  PerlIO_binmode(pTHX_ PerlIO *f, int ptype, int imode,
                  const char *layers);
void  PerlIO_debug(const char *fmt,...);

```

## DESCRIPTION

Perl's source code, and extensions that want maximum portability, should use the above functions instead of those defined in ANSI C's *stdio.h*. The perl headers (in particular "perlio.h") will "#define" them to the I/O mechanism selected at Configure time.

The functions are modeled on those in *stdio.h*, but parameter order has been "tidied up a little".

"PerlIO \*" takes the place of FILE \*. Like FILE \* it should be treated as opaque (it is probably safe to assume it is a pointer to something).

There are currently two implementations:

#### 1. USE\_STDIO

All above are #define'd to stdio functions or are trivial wrapper functions which call stdio. In this case *only* PerlIO \* is a FILE \*. This has been the default implementation since the abstraction was introduced in perl5.003\_02.

#### 2. USE\_PERLIO

Introduced just after perl5.7.0, this is a re-implementation of the above abstraction which allows perl more control over how IO is done as it decouples IO from the way the operating system and C library choose to do things. For USE\_PERLIO PerlIO \* has an extra layer of indirection - it is a pointer-to-a-pointer. This allows the PerlIO \* to remain with a known value while swapping the implementation around underneath *at run time*. In this case all the above are true (but very simple) functions which call the underlying implementation.

This is the only implementation for which "PerlIO\_apply\_layers()" does anything "interesting".

The USE\_PERLIO implementation is described in perliol.

Because "perlio.h" is a thin layer (for efficiency) the semantics of these functions are somewhat dependent on the underlying implementation. Where these variations are understood they are noted below.

Unless otherwise noted, functions return 0 on success, or a negative value (usually "EOF" which is usually -1) and set "errno" on error.

#### **PerlIO\_stdin(), PerlIO\_stdout(), PerlIO\_stderr()**

Use these rather than "stdin", "stdout", "stderr". They are written to look like "function calls" rather than variables because this makes it easier to *make them* function calls if platform cannot export data to loaded modules, or if (say) different "threads" might have different values.

#### **PerlIO\_open(path, mode), PerlIO\_fdopen(fd,mode)**

These correspond to **fopen()/fdopen()** and the arguments are the same. Return "NULL" and set "errno" if there is an error. There may be an implementation limit on the number of open handles, which may be lower than the limit on the number of open files - "errno" may not be set when "NULL" is returned if this limit is exceeded.

#### **PerlIO\_reopen(path,mode,f)**

While this currently exists in both implementations, perl itself does not use it. *As perl does not use it, it is not well tested.*

Perl prefers to "dup" the new low-level descriptor to the descriptor used by the existing PerlIO. This may become the behaviour of this function in the future.

**PerlIO\_printf(f,fmt,...), PerlIO\_vprintf(f,fmt,a)**

These are **fprintf()/vfprintf()** equivalents.

**PerlIO\_stdoutf(fmt,...)**

This is **printf()** equivalent. `printf` is #defined to this function, so it is (currently) legal to use `"printf(fmt,...)"` in perl sources.

**PerlIO\_read(f,buf,count), PerlIO\_write(f,buf,count)**

These correspond functionally to **fread()** and **fwrite()** but the arguments and return values are different. The **PerlIO\_read()** and **PerlIO\_write()** signatures have been modeled on the more sane low level **read()** and **write()** functions instead: The "file" argument is passed first, there is only one "count", and the return value can distinguish between error and "EOF".

Returns a byte count if successful (which may be zero or positive), returns negative value and sets "errno" on error. Depending on implementation "errno" may be "EINTR" if operation was interrupted by a signal.

**PerlIO\_fill(f)**

Fills the buffer associated with "f" with data from the layer below. "PerlIO\_read" calls this as part of its normal operation. Returns 0 upon success; -1 on failure.

**PerlIO\_close(f)**

Depending on implementation "errno" may be "EINTR" if operation was interrupted by a signal.

**PerlIO\_puts(f,s), PerlIO\_putc(f,c)**

These correspond to **fputs()** and **fputc()**. Note that arguments have been revised to have "file" first.

**PerlIO\_ungetc(f,c)**

This corresponds to **ungetc()**. Note that arguments have been revised to have "file" first. Arranges that next read operation will return the byte **c**. Despite the implied "character" in the name only values in the range 0..0xFF are defined. Returns the byte **c** on success or -1 ("EOF") on error. The number of bytes that can be "pushed back" may vary, only 1 character is certain, and then only if it is the last character that was read from the handle.

**PerlIO\_unread(f,buf,count)**

This allows one to unget more than a single byte. It effectively unshifts "count" bytes onto the

beginning of the buffer "buf", so that the next read operation(s) will return them before anything else that was in the buffer.

Returns the number of unread bytes.

### **PerlIO\_getc(f)**

This corresponds to **getc()**. Despite the c in the name only byte range 0..0xFF is supported. Returns the character read or -1 ("EOF") on error.

### **PerlIO\_eof(f)**

This corresponds to **feof()**. Returns a true/false indication of whether the handle is at end of file. For terminal devices this may or may not be "sticky" depending on the implementation. The flag is cleared by **PerlIO\_seek()**, or **PerlIO\_rewind()**.

### **PerlIO\_error(f)**

This corresponds to **ferror()**. Returns a true/false indication of whether there has been an IO error on the handle.

### **PerlIO\_fileno(f)**

This corresponds to **fileno()**, note that on some platforms, the meaning of "fileno" may not match Unix. Returns -1 if the handle has no open descriptor associated with it.

### **PerlIO\_clearerr(f)**

This corresponds to **clearerr()**, i.e., clears 'error' and (usually) 'eof' flags for the "stream". Does not return a value.

### **PerlIO\_flush(f)**

This corresponds to **fflush()**. Sends any buffered write data to the underlying file. If called with "NULL" this may flush all open streams (or core dump with some USE\_STDIO implementations). Calling on a handle open for read only, or on which last operation was a read of some kind may lead to undefined behaviour on some USE\_STDIO implementations. The USE\_PERLIO (layers) implementation tries to behave better: it flushes all open streams when passed "NULL", and attempts to retain data on read streams either in the buffer or by seeking the handle to the current logical position.

### **PerlIO\_seek(f,offset,whence)**

This corresponds to **fseek()**. Sends buffered write data to the underlying file, or discards any buffered read data, then positions the file descriptor as specified by **offset** and **whence** (sic). This is the correct thing to do when switching between read and write on the same handle (see issues with **PerlIO\_flush()** above). Offset is of type "Off\_t" which is a perl Configure value which may

not be same as `stdio`'s "off\_t".

### **PerlIO\_tell(f)**

This corresponds to **ftell()**. Returns the current file position, or (Off\_t) -1 on error. May just return value system "knows" without making a system call or checking the underlying file descriptor (so use on shared file descriptors is not safe without a **PerlIO\_seek()**). Return value is of type "Off\_t" which is a perl Configure value which may not be same as `stdio`'s "off\_t".

### **PerlIO\_getpos(f,p), PerlIO\_setpos(f,p)**

These correspond (loosely) to **fgetpos()** and **fsetpos()**. Rather than `stdio`'s `Fpos_t` they expect a "Perl Scalar Value" to be passed. What is stored there should be considered opaque. The layout of the data may vary from handle to handle. When not using `stdio` or if platform does not have the `stdio` calls then they are implemented in terms of **PerlIO\_tell()** and **PerlIO\_seek()**.

### **PerlIO\_rewind(f)**

This corresponds to **rewind()**. It is usually defined as being

```
PerlIO_seek(f,(Off_t)0L, SEEK_SET);
PerlIO_clearerr(f);
```

### **PerlIO\_tmpfile()**

This corresponds to **tmpfile()**, i.e., returns an anonymous PerlIO or NULL on error. The system will attempt to automatically delete the file when closed. On Unix the file is usually "unlink"-ed just after it is created so it does not matter how it gets closed. On other systems the file may only be deleted if closed via **PerlIO\_close()** and/or the program exits via "exit". Depending on the implementation there may be "race conditions" which allow other processes access to the file, though in general it will be safer in this regard than ad. hoc. schemes.

### **PerlIO\_setlinebuf(f)**

This corresponds to **setlinebuf()**. Does not return a value. What constitutes a "line" is implementation dependent but usually means that writing "\n" flushes the buffer. What happens with things like "this\nthat" is uncertain. (Perl core uses it *only* when "dumping"; it has nothing to do with \$| auto-flush.)

## **Co-existence with stdio**

There is outline support for co-existence of PerlIO with `stdio`. Obviously if PerlIO is implemented in terms of `stdio` there is no problem. However in other cases then mechanisms must exist to create a FILE \* which can be passed to library code which is going to use `stdio` calls.

The first step is to add this line:

```
#define PERLIO_NOT_STDIO 0
```

*before* including any perl header files. (This will probably become the default at some point). That prevents "perlio.h" from attempting to #define stdio functions onto PerlIO functions.

XS code is probably better using "typemap" if it expects FILE \* arguments. The standard typemap will be adjusted to comprehend any changes in this area.

### **PerlIO\_importFILE(f,mode)**

Used to get a PerlIO \* from a FILE \*.

The mode argument should be a string as would be passed to fopen/PerlIO\_open. If it is NULL then - for legacy support - the code will (depending upon the platform and the implementation) either attempt to empirically determine the mode in which *f* is open, or use "r+" to indicate a read/write stream.

Once called the FILE \* should *ONLY* be closed by calling "PerlIO\_close()" on the returned PerlIO \*.

The PerlIO is set to textmode. Use PerlIO\_binmode if this is not the desired mode.

This is **not** the reverse of **PerlIO\_exportFILE()**.

### **PerlIO\_exportFILE(f,mode)**

Given a PerlIO \* create a 'native' FILE \* suitable for passing to code expecting to be compiled and linked with ANSI C *stdio.h*. The mode argument should be a string as would be passed to fopen/PerlIO\_open. If it is NULL then - for legacy support - the FILE \* is opened in same mode as the PerlIO \*.

The fact that such a FILE \* has been 'exported' is recorded, (normally by pushing a new :stdio "layer" onto the PerlIO \*), which may affect future PerlIO operations on the original PerlIO \*. You should not call "fclose()" on the file unless you call "PerlIO\_releaseFILE()" to disassociate it from the PerlIO \*. (Do not use **PerlIO\_importFILE()** for doing the disassociation.)

Calling this function repeatedly will create a FILE \* on each call (and will push an :stdio layer each time as well).

### **PerlIO\_releaseFILE(p,f)**

Calling PerlIO\_releaseFILE informs PerlIO that all use of FILE \* is complete. It is removed from the list of 'exported' FILE \*s, and the associated PerlIO \* should revert to its original behaviour.

Use this to disassociate a file from a PerlIO \* that was associated using **PerlIO\_exportFILE()**.

### **PerlIO\_findFILE(f)**

Returns a native FILE \* used by a stdio layer. If there is none, it will create one with PerlIO\_exportFILE. In either case the FILE \* should be considered as belonging to PerlIO subsystem and should only be closed by calling "PerlIO\_close()".

### **"Fast gets" Functions**

In addition to standard-like API defined so far above there is an "implementation" interface which allows perl to get at internals of PerlIO. The following calls correspond to the various FILE\_xxx macros determined by Configure - or their equivalent in other implementations. This section is really of interest to only those concerned with detailed perl-core behaviour, implementing a PerlIO mapping or writing code which can make use of the "read ahead" that has been done by the IO system in the same way perl does. Note that any code that uses these interfaces must be prepared to do things the traditional way if a handle does not support them.

### **PerlIO\_fast\_gets(f)**

Returns true if implementation has all the interfaces required to allow perl's "sv\_gets" to "bypass" normal IO mechanism. This can vary from handle to handle.

```
PerlIO_fast_gets(f) = PerlIO_has_cntptr(f) && \
    PerlIO_canset_cnt(f) && \
    'Can set pointer into buffer'
```

### **PerlIO\_has\_cntptr(f)**

Implementation can return pointer to current position in the "buffer" and a count of bytes available in the buffer. Do not use this - use PerlIO\_fast\_gets.

### **PerlIO\_get\_cnt(f)**

Return count of readable bytes in the buffer. Zero or negative return means no more bytes available.

### **PerlIO\_get\_ptr(f)**

Return pointer to next readable byte in buffer, accessing via the pointer (dereferencing) is only safe if **PerlIO\_get\_cnt()** has returned a positive value. Only positive offsets up to value returned by **PerlIO\_get\_cnt()** are allowed.

### **PerlIO\_set\_ptrcnt(f,p,c)**

Set pointer into buffer, and a count of bytes still in the buffer. Should be used only to set pointer to within range implied by previous calls to "PerlIO\_get\_ptr" and "PerlIO\_get\_cnt". The two values



*must* be consistent with each other (implementation may only use one or the other or may require both).

### **PerlIO\_canset\_cnt(f)**

Implementation can adjust its idea of number of bytes in the buffer. Do not use this - use `PerlIO_fast_gets`.

### **PerlIO\_set\_cnt(f,c)**

Obscure - set count of bytes in the buffer. Deprecated. Only usable if `PerlIO_canset_cnt()` returns true. Currently used in only `doio.c` to force count less than -1 to -1. Perhaps should be `PerlIO_set_empty` or similar. This call may actually do nothing if "count" is deduced from pointer and a "limit". Do not use this - use `PerlIO_set_ptrcnt()`.

### **PerlIO\_has\_base(f)**

Returns true if implementation has a buffer, and can return pointer to whole buffer and its size. Used by perl for **-T** / **-B** tests. Other uses would be very obscure...

### **PerlIO\_get\_base(f)**

Return *start* of buffer. Access only positive offsets in the buffer up to the value returned by `PerlIO_get_bufsiz()`.

### **PerlIO\_get\_bufsiz(f)**

Return the *total number of bytes* in the buffer, this is neither the number that can be read, nor the amount of memory allocated to the buffer. Rather it is what the operating system and/or implementation happened to "read()" (or whatever) last time IO was requested.

## **Other Functions**

`PerlIO_apply_layers(aTHX_ f,mode,layers)`

The new interface to the USE\_PERLIO implementation. The layers `":crlf"` and `":raw"` are the only ones allowed for other implementations and those are silently ignored. (As of perl5.8 `":raw"` is deprecated.) Use `PerlIO_binmode()` below for the portable case.

`PerlIO_binmode(aTHX_ f,ptype,imode,layers)`

The hook used by perl's "binmode" operator. **ptype** is perl's character for the kind of IO:

'<' read

'>' write

'+' read/write

**imode** is `"O_BINARY"` or `"O_TEXT"`.

**layers** is a string of layers to apply; only ":crlf" makes sense in the non-USE\_PERLIO case. (As of perl5.8 ":raw" is deprecated in favour of passing NULL.)

Portable cases are:

```
PerlIO_binmode(aTHX_ f,ptype,O_BINARY,NULL);
and
PerlIO_binmode(aTHX_ f,ptype,O_TEXT,":crlf");
```

On Unix these calls probably have no effect whatsoever. Elsewhere they alter "\n" to CR,LF translation and possibly cause a special text "end of file" indicator to be written or honoured on read. The effect of making the call after doing any IO to the handle depends on the implementation. (It may be ignored, affect any data which is already buffered as well, or only apply to subsequent data.)

PerlIO\_debug(fmt,...)

PerlIO\_debug is a **printf()**-like function which can be used for debugging. No return value. Its main use is inside PerlIO where using real printf, **warn()** etc. would recursively call PerlIO and be a problem.

PerlIO\_debug writes to the file named by \$ENV{'PERLIO\_DEBUG'} or defaults to stderr if the environment variable is not defined. Typical use might be

Bourne shells (sh, ksh, bash, zsh, ash, ...):

```
PERLIO_DEBUG=/tmp/perliodebug.log ./perl -Di somescript some args
```

Csh/Tcsh:

```
setenv PERLIO_DEBUG /tmp/perliodebug.log
./perl -Di somescript some args
```

If you have the "env" utility:

```
env PERLIO_DEBUG=/tmp/perliodebug.log ./perl -Di somescript args
```

Win32:

```
set PERLIO_DEBUG=perliodebug.log
perl -Di somescript some args
```

On a Perl built without "-DDEBUGGING", or when the "-Di" command-line switch is not specified, or under taint, **PerlIO\_debug()** is a no-op.