

63.2.2 The *poll()* System Call

The *poll()* system call performs a similar task to *select()*. The major difference between the two system calls lies in how we specify the file descriptors to be monitored. With *select()*, we provide three sets, each marked to indicate the file descriptors of interest. With *poll()*, we provide a list of file descriptors, each marked with the set of events of interest.

```
#include <poll.h>

int poll(struct pollfd fds[], nfds_t nfds, int timeout);

    Returns number of ready file descriptors, 0 on timeout, or -1 on error
```

The *fds* argument and the *pollfd* array (*nfds*) specify the file descriptors that *poll()* is to monitor. The *timeout* argument can be used to set an upper limit on the time for which *poll()* will block. We describe each of these arguments in detail below.

The *pollfd* array

The *fds* argument lists the file descriptors to be monitored by *poll()*. This argument is an array of *pollfd* structures, defined as follows:

```
struct pollfd {
    int    fd;           /* File descriptor */
    short  events;        /* Requested events bit mask */
    short  revents;       /* Returned events bit mask */
};
```

The *nfds* argument specifies the number of items in the *fds* array. The *nfds_t* data type used to type the *nfds* argument is an unsigned integer type.

The *events* and *revents* fields of the *pollfd* structure are bit masks. The caller initializes *events* to specify the events to be monitored for the file descriptor *fd*. Upon return from *poll()*, *revents* is set to indicate which of those events actually occurred for this file descriptor.

Table 63-2 lists the bits that may appear in the *events* and *revents* fields. The first group of bits in this table (POLLIN, POLLRDNORM, POLLRDBAND, POLLPRI, and POLLRDHUP) are concerned with input events. The next group of bits (POLLOUT, POLLWRNORM, and POLLWRBAND) are concerned with output events. The third group of bits (POLLERR, POLLHUP, and POLLNVAL) are set in the *revents* field to return additional information about the file descriptor. If specified in the *events* field, these three bits are ignored. The final bit (POLLMSG) is unused by *poll()* on Linux.

On UNIX implementations providing STREAMS devices, POLLMSG indicates that a message containing a SIGPOLL signal has reached the head of the stream. POLLMSG is unused on Linux, because Linux doesn't implement STREAMS.