FlexProp Reference

5.4.3

Total Spectrum Software

05/08/2021

Contents

General Compiler Features

This document describes compiler features common to all of the Flex languages.

Languages

The compiler supports Spin (in both Spin 1 and Spin 2 variants), C (and a subset of C++), and BASIC. The choice of which language to use is determined by the file extension.

Spin OBJ blocks

In Spin OBJ blocks, the default extension for file names is the same as the file name of the source file; that is:

```
OBJ
  a: "somedevice"
```

will look for "somedevice.spin2" first in a .spin2 file, then "somedevice.spin"; in a .spin file the order is reversed.

In order to avoid ambiguity it is suggested that the file name extension always be explicitly given. An explicit file name is always required for non-Spin objects:

```
OBJ
  a: "somedevice.spin"
  b: "otherdevice.bas"
```

Objects in other languages

In BASIC `class using` and C `struct __using` the full filename of the source object, including extension, must always be used.

BASIC and C also allow inline declarations of classes, using `class`. See the respective language documents for more details.

Fast Cache (Fcache)

Fcache is a special feature of the compiler whereby small loops are copied from HUB memory into local (COG) memory before execution. This speeds up repated loops quite a bit. Fcache is available only if optimization is enabled.

Some inline assembly blocks may also be marked to be copied to fcache before execution; see the section on inline assembly for a description of this.

What loops will be placed in fcache

Loops will be placed in fcache only if (a) they will fit, and (b) they contain no branches to outside the loop (including subroutine calls). The size of the fcache may be set by the `--fcache` flag, but is generally 1024 bytes on P2 and 128 bytes on P1.

Inline assembly

All of the languages allow inline assembly within functions. There are 3 different forms of inline assembly:

(1) Plain inline assembly. This is generated by `asm`/`endasm` in Spin and BASIC, and from an `__asm { }` block in C. These blocks run in hubexec mode (for P2) or LMM (for P1) and are optimized by the optimizer.

(2) HUB non-optimized assembly. This is generated by `asm const`/`end asm` in BASIC and `__asm const{}` in C; it is not currently available in Spin. Like plain assembly this runs from HUB, but is not subject to optimization.

(3) FCACHEd non-optimized assembly. This is generated by `org`/`end` in Spin, `asm cpu`/`end` in BASIC, and `__asm volatile{}` in C. This is not subject to optimization, and before execution it is loaded into the FCACHE area, so its timing is based on running from internal memory rather than HUB.

Restrictions in inline assembly

Inline assembly within a function follows a different path through the compiler than "regular" assembly in a DAT section (Spin) or `asm shared` (BASIC). This has a number of consequences; not all constructions will work properly, and the inline assembly can be limited.

Only local variables

Only hardware registers and variables local to the function may be used in inline assembly. Global or method variables may not be referenced.

Local variables not usable in some functions

If a function takes the address of a parameter, or of a local variable, then its local variables are placed on the stack and may not be referred to in inline assembly.

No branches outside the function

Branching within a function is supported in inline assembly, but trying to branch outside the function or to call another function is not supported. The results are undefined; calls in particular may appear to work in some cases, but then fail when the called function is modified.

It is also not legal to return from inside inline assembly. For Spin2, returns are automatically converted to jumps to the end of the inline assembly (this is for compatibility with PNut), but it's probably better to write this explicitly yourself.

No register declarations

Do not try to declare registers; the inline assembly probably will not be running from COG memory. If you need some scratch registers in inline assembly, declare them as local variables in the function.

General Guidelines

Try to keep inline assembly as simple as possible. Use the high level language for loops and conditional control structures; the high level language is there for a reason!

Functions in COG or LUT memory

Normally functions are placed in HUB memory, because there is a lot more of that. However, it is possible to force some functions to be placed in the chip's internal memory, where they will execute much more quickly. This must be done with care, because internal memory is a very limited resource.

Only small functions should be placed in internal memory, and these functions should not call any other function.

Spin/Spin2

To put a method into COG memory, place a special comment `{++cog}` after the `PUB` or `PRI` declaration of the method.

```
pub {++cog} add(x, y)
  return x+y
```

Similarly, to put the method into LUT memory (on the P2 only, obviously) then use the comment `{++lut}`.

BASIC

Place the keyword `for` before the function or subroutine's name in its declaration, followed by a string specifying the memory (`"cog"` or `"lut"`):

```
function for "cog" toupper(c as ubyte) as ubyte
  if c >= asc("a") and c <= asc("z") then
    c = c + (asc("A") - asc("a"))
  end if
  return c
end function
```

C/C++

Place `__attribute__((cog))` after the function declaration but before its body:

```
int add(int x, int y) __attribute__((cog))
{
  return x+y;
}
```

Similarly use `__attribute__((lut))` to place the function into LUT memory.

Register Usage

P1

Pretty much all of COG RAM is used by the compiler. No specific hardware registers are used.

P2

Most of COG RAM is used by the compiler, except that $1e0-$1ef are left free for application use. COG RAM from $00 to $ff is used for FCACHE, and so when you are sure no FCACHE is in use you may use this for scratch.

The second half of LUT memory (from $300 to $3ff) may be used by compiler internal functions.

`ptra` is used for the stack pointer. Applications should avoid using it.

`pa` is used internally for fcache loading. Applications may use it as a temporary variable, but be aware that any code execution which may trigger an fcache load (e.g. any loop or subroutine call) may trash its value.

Memory Map

HUB

Code starts at 0 in HUB (by default, there are command line options to change this). Data starts after the code. The heap is part of the data area. The stack starts after this and grows upwards.

COG

Most of COG RAM is used by the compiler, except that $1e0-$1ef is left free for application use.

LUT

The first part of LUT memory (from $200 to $300) is used for any functions explicitly placed into LUT. The LUT memory from $300 to $400 (the second half of LUT) is used for internal purposes.

Optimizations

Listed below are optimizations which may be enabled on the command line or on a per-function basis.

Multiplication conversion (always)

Multiplies by powers of two, or numbers near a power of two, are converted to shifts. For example

```
a := a*10
```

is converted to

```
a := (a<<3) + (a<<1)
```

A similar optimization is performed for divisions by powers of two.

Unused method removal (-O1, -Oremove-unused)

This is pretty standard; if a method is not used, no code is emitted for it.

Dead code elimination (-O1, -Oremove-dead)

Within functions if code can obviously never be reached it is also removed. So for instance in something like:

```
CON
  pin = 1
...
if (pin == 2)
  foo
```

The if statement and call to `foo` are removed since the condition is always false.

Small Method inlining (-O1, -Oinline-small)

Very small methods are expanded inline. This may be prevented by declaring the method with the "noinline" attribute.

Register optimization (-O1, -Oregs)

The compiler analyzes assignments to registers and attempts to minimize the number of moves (and temporary registers) required.

Branch elimination (-O1, -Obranch-convert)

Short branch sequences are converted to conditional execution where possible.

Constant propagation (-O1, -Oconst)

If a register is known to contain a constant, arithmetic on that register can often be replaced with move of another constant.

Peephole optimization (-O1, -Opeephole)

In generated assembly code, various shorter combinations of instructions can sometimes be substituted for longer combinations.

Tail call optimization (-O1, -Otail-calls)

Convert recursive calls into jumps when possible

Loop optimization (-O1, -Oloop-basic)

In some circumstances the optimizer can re-arrange counting loops so that the `djnz` instruction may be used instead of a combination of add/sub, compare, and branch. In -O2 a more thorough loop analysis makes this possible in more cases.

Fcache (-O1, -Ofcache)

Small loops are copied to internal memory (COG) to be executed there. These loops cannot have any non-inlined calls in them.

Single Use Method inlining (-O2, -Oinline-single)

If a method is called only once in a whole program, it is expanded inline at the call site.

Common Subexpression Elimination (-O2, -Ocse)

Code like:

```
c := a*a + a*a
```

is automaticaly converted to something like:

```
tmp := a*a
c := tmp + tmp
```

Loop Strength Reduction (-O2, -Oloop-reduce)

Array indexes

Array lookups inside loops are converted to pointers. So:

```
repeat i from 0 to n-1
   a[i] := b[i]
```

is converted to the equivalent of

```
aptr := @a[0]
bptr := @b[0]
repeat n
  long[aptr] := long[bptr]
  aptr += 4
  bptr += 4
```

Multiply to addition

An expression like `(i*100)` where `i` is a loop index can be converted to something like `itmp \ itmp + 100`

Per-function control of optimizations

It is possible to enable or disable individual optimizations in a function by using attributes. For example, to disable loop reduction for a particular C function, one would add an attribute:

```
int foo(int x) __attribute__(opt(!loop-reduce)) {
...
}
```

A similar effect is achieved in Spin by adding a comment `{++opt(!loop-reduce)}` between the `pub` or `pri` and the function name.

In BASIC we use the `for` keyword followed by a string giving the optimization options:

```
function for "opt(!loop-reduce)" myfunc()
```

Multiple options may be given, separated by commas. To turn an option off, prefix it with `!` or with `~`. To enable all options for a particular optimization level, start the string with `0`, `1`, `2`, etc., or with the word `all` to enable all optimizations (regardless of the compiler optimization level chosen).

Thus, a Spin function with `{++opt(0,peephole)}` will always be compiled with no optimization except peepholes, even when the `-O2` option is given to the compiler.

Optimization control on the command line

Multiple `-O` options may be given, or combined separated by commas. So for example to compile with no optimizations except basic register and peephole, one would give `-O0,regs,peephole`. To compile with `-O2` but with peepholes turned off, one would give `-O2,!peephole`.

Memory Management

There are some built in functions for doing memory allocation. These are intended for C or BASIC, but may be used by Spin programs as well.

Heap allocation

The main function is `_gc_alloc_managed(siz)`, which allocates `siz` bytes of memory managed by the garbage collector. It returns 0 if not enough memory is avilable, otherwise returns a pointer to the start of the memory (like C's `malloc`). As long as there is some reference in COG or HUB memory to the pointer which got returned, the memory will be considered "in use". If there is no more such reference then the garbage collector will feel free to reclaim it. There's also `_gc_alloc(siz)` which is similar but marks the memory so it will never be reclaimed, and `_gc_free(ptr)` which explicitly frees a pointer previously allocated by `_gc_alloc` or `_gc_alloc_managed`.

The size of the heap is determined by a constant `HEAPSIZE` declared in the top level object. If none is given then a (small) default value is used.

Example:

```
' put this CON in the top level object to specify how much memory should be provided for
' memory allocation (the "heap"). The default is 4K on P2, 256 bytes on P1
CON
   HEAPSIZE = 32768 ' or however much memory you want to provide for the allocator

' here's a function to allocate memory
' "siz" is the size in bytes
PUB allocmem(size) : ptr
  ptr := _gc_alloc_managed(size)
```

The garbage collection functions and heap are only included in programs which explicitly ask for them.

Heap size specification

In SPIN:

```
CON HEAPSIZE=32768
```

In BASIC:

```
const HEAPSIZE=32768
```

In C:

```
enum { HEAPSIZE=32768 };
```

The C version is a little unexpected; one would expect HEAPSIZE to be declared as `const int` or with `#define`. This is a technical limitation that I hope to fix someday.

Stack allocation

Temporary memory may be allocated on the stack by means of the call `__builtin_alloca(siz)`, which allocates `siz` bytes of memory on the stack. This is like the C `alloca` function. Note that the pointer returned by `__builtin_alloca` will become invalid as soon as the current function returns, so it should not be placed in any global variable (and definitely should not be returned from the function!)

Terminal Control

The FlexProp system uses the default system terminal, configured when possible to accept VT100 (ANSI) escape sequences.

Changing baud rate

All languages have a `_setbaud(N)` function to set the baud rate to `N`.

Changing echo and CR/LF interpretation

Normally input (e.g. from a C `getchar()`) is echoed back to the screen, and carriage return (ASCII 13) is converted to line feed (ASCII 10). Both of these behaviors may be changed via the `_setrxtxflags(mode)` function. The bits in

`mode` control terminal behavior: if `mode & 1` is true, then characters are echoed, and if `mode & 2` is true then carriage return (CR) is converted to line feed (LF) on input, and line feed is converted to CR + LF on output.

The current state of the flags may be retrieved via `_getrxtxflags()`.

File I/O (P2 Only)

C and BASIC have built in support for accessing file systems. The file systems first must be given a name with the `mount` system call, and then may be accessed with the normal language functions.

Mount

The `mount` call gives a name to a file system. For example, after

```
mount("/host", _vfs_open_host());
mount("/sd", _vfs_open_sdcard());
```

files on the host PC may be accessed via names like "/host/foo.txt", "/host/bar/bar.txt", and so on, and files on the SD card may be accessed by names like "/sd/root.txt", "/sd/subdir/file.txt", and so on.

This only works on P2, because it requires a lot of HUB memory. Also, the host file server requires features built in to `loadp2`.

Available file systems are:

- `_vfs_open_host()` (for the loadp2 Plan 9 file system)
- `_vfs_open_sdcard()` for a FAT file system on the P2 SD card (using default pins 58-61)
- `_vfs_open_sdcardx()` for a FAT file system on SD card using custom pins

It is OK to make multiple mount calls, but they should have different names.

Options for SD Card

If you define the symbol `FF_USE_LFN` on the command line with an option like `-DFF_USE_LFN` then long file names will be enabled for the SD card.

The pins to use for the SD card may be changed by using `_vfs_open_sdcardx` instead of `_vfs_open_sdcard`. The parameters for `_vfs_open_sdcardx` are the clock pin, select pin, data in, and data out pins, in that order. Thus, `_vfs_open_sdcard` is actually equivalent to `_vfs_open_sdcardx(61, 60, 59, 58)`.

Command Line Options

Options for flexspin

There are various command line options for the compiler which may modify the compilation:

```
  [ --version ]      print just the compiler version, then exit
  [ -h ]             display this help
  [ -L or -I <path> ] add a directory to the include path
  [ -o ]             output filename
  [ -b ]             output binary file format
  [ -e ]             output eeprom file format
  [ -c ]             output only DAT sections
  [ -l ]             output a .lst listing file
  [ -f ]             output list of file names
  [ -g ]             enable debug statements
  [ -q ]             quiet mode (suppress banner and non-error text)
  [ -p ]             disable the preprocessor
  [ -O[#] ]          set optimization level
                       -O0 disable all optimization
                       -O1 apply default optimization (same as no -O flag)
              -O2 apply all optimization (same as -O)
  [ -Wall ]          enable all warnings, including warnings about language extensions
  [ -Werror ]        turn warnings into errors
  [ -Wabs-paths ]    print absolute paths for file names in errors/warnings
  [ -Wmax-errors=N ] allow at most N errors in a pass before stopping
  [ -D <define> ]    add a define
  [ -2 ]             compile for Prop2
  [ -w ]             produce Spin wrappers for PASM code
  [ -H nnnn ]        change the base HUB address (see below)
  [ -E ]             omit any coginit header
  [ --code=cog  ]    compile to run in COG memory instead of HUB
  [ --fcache=N  ]    set size of FCACHE space in longs (0 to disable)
  [ --fixedreal ]    use 16.16 fixed point instead of IEEE floating point
  [ --lmm=xxx ]      use alternate LMM implementation for P1
           xxx = orig uses original flexspin LMM
           xxx = slow uses traditional (slow) LMM
  [ --tabs=N ]       specifify number of spaces between tab stops (default 8)
```

`flexspin.exe` checks the name it was invoked by. If the name starts with the string "bstc" (case matters) then its output messages mimic that of the bstc compiler; otherwise it tries to match openspin's messages. This is for compatibility with Propeller IDE. For example, you can use flexspin with the PropellerIDE by renaming `bstc.exe` to `bstc.orig.exe` and then copying `flexspin.exe` to `bstc.exe`.

Options for flexcc

`flexcc` is similar to `flexspin`, but has arguments more like the traditional `cc` command line compiler.

```
[ --help ]          display this help
[ -c ]              output only .o file
[ -D <define> ]     add a define
[ -g ]              include debug info in output
[ -L or -I <path> ] add a directory to the include path
[ -o <name> ]       set output filename to <name>
[ -2 ]              compile for Prop2
[ -O# ]             set optimization level:
        -O0 = no optimization
        -O1 = basic optimization
        -O2 = all optimization
[ -Wall ]           enable warnings for language extensions and other features
[ -Werror ]         make warnings into errors
[ -Wabs-paths ]     print absolute paths for file names in errors/warnings
[ -Wmax-errors=N ] allow at most N errors in a pass before stopping
[ -x ]              capture program exit code (for testing)
[ --code=cog ]      compile for COG mode instead of LMM
[ --fcache=N ]      set FCACHE size to N (0 to disable)
[ --fixedreal ]     use 16.16 fixed point in place of floats
[ --lmm=xxx ]       use alternate LMM implementation for P1
         xxx = orig uses original flexspin LMM
         xxx = slow uses traditional (slow) LMM
[ --version ]       just show compiler version
```

Changing Hub address

In P2 mode, you may want to change the base hub address for the binary. Normally P2 binaries start at the standard offset of `0x400`. But if you want, for example, to load a flexspin compiled program from TAQOZ or some similar program, you may want to start at a different address (TAQOZ uses the first 64K of RAM). To do this, you may use some combination of the `-H` and `-E` flags.

`-H nnnn` changes the base HUB address from `0x400` to `nnnn`, where `nnnn` is either a decimal number like `65536` or a hex number prefixed with `0x`. By default the binary still expects to be loaded at address 0, so it starts with a `coginit #0, ##nnnn` instruction and then zero padding until the hub start. To skip the `coginit` and padding, add the `-E` flag.

Example

To compile a program to start at address 65536 (at the 64K boundary), do:

```
flexspin -2 -H 0x10000 -E fibo.bas
```

Common low level functions

A number of low level functions are available in all languages. The C prototypes are given below, but they may be called from any language and are always available. If a user function with the same name is provided, the built-in function will not be available from user code (but internally the libraries *may* continue to use the built-in version; this isn't defined).

Unless otherwise noted, these functions are available for both P1 and P2.

Serial port access

__txraw

```
int _txraw(int c)
```

sends character `c` out the default serial port. Always returns 1.

__rxraw

```
int _rxraw(int n=0)
```

Receives a character on the default serial port. `n` is a timeout in milliseconds. If the timeout elapses with no character received, `_rxraw` returns -1, otherwise it returns the received character. The default timeout (0) signifies "forever"; that is, if `n` is 0 the `_rxraw` function will wait as long as necessary until a character is received.

__setbaud

```
void _setbaud(int rate)
```

Sets the baud rate on the default serial port to `rate`. For example, to change the serial port to 115200 baud you would call `_setbaud(115200)`. The default rates set up in C and BASIC initialization code are are 115200 for P1 and 230400 for P2. In Spin you may need to call `_setbaud` explicitly before calling `_rxraw` or `_txraw`.

Time related functions

__getsec

Gets elapsed seconds since the system was booted. Uses the system clock, which wraps around after about 50 seconds on the P1.

__getms

Gets elapsed milliseconds since boot.

__getus

Gets elapsed microseconds since boot.

__waitx

```
void _waitx(unsigned cycles)
```

Pauses for `cycles` cycles. Note that the maximum waiting period is about half of the system clock frequency.

__waitms

```
void _waitms(unsigned ms)
```

Wait for `ms` milliseconds. For waits of more than a second, this function will loop internally so as to avoid limits of the 32 bit clock counter.

__waitus

```
void _waitus(unsigned us)
```

Wait for `us` microseconds. For waits of more than a second, this function will loop internally so as to avoid limits of the 32 bit clock counter.

Cog control

__cogchk

```
int _cogchk(int id)
```

Checks to see if cog number id is running. Returns -1 if running, 0 if not.

This function may be relatively slow on P1, as it has to manually probe the COGs (on P2 it's a built in instruction).