dyn-call

Version 0.7

Daniel Adler (dadler@uni-goettingen.de)
Tassilo Philipp (tphilipp@potion-studios.com)

January 2, 2012
## Contents

1 Motivation 5
   1.1 Static function calls in C 5
   1.2 Anatomy of machine-level calls 5

2 Overview 7
   2.1 Features 7
   2.2 Showcase 8
   2.3 Supported platforms/architectures 9

3 Building the library 10
   3.1 Requirements 10
   3.2 Supported/tested platforms and build systems 10
   3.3 Build instructions 12
   3.4 Build-tool specific notes 13
   3.5 Build with CMake 13

4 Bindings to programming languages 14
   4.1 Common Architecture 14
      4.1.1 Dynamic loading of code 14
      4.1.2 Functions 14
      4.1.3 Signatures 15
   4.2 Python language bindings 16
   4.3 R language bindings 16
   4.4 Ruby language bindings 17

5 Library Design 18
   5.1 Design considerations 18

6 Developers 19
   6.1 Project root 19
   6.2 Test suites 19

7 Epilog 21
   7.1 Stability and security considerations 21
   7.2 Embedding 21
   7.3 Multi-threading 21
   7.4 Supported types 21
   7.5 Roadmap 21
   7.6 Related libraries 21
A  Dynccall C library API
   A.1  Supported C/C++ argument and return types ........................................ 22
   A.2  Call Virtual Machine - CallVM ............................................................ 23
   A.3  Allocation ................................................................................................. 23
   A.4  Error Reporting ......................................................................................... 23
   A.5  Configuration ............................................................................................. 24
   A.6  Machine state reset .................................................................................... 25
   A.7  Argument binding ....................................................................................... 25
   A.8  Call invocation ............................................................................................ 25
   A.9  Formatted calls (ANSI C ellipsis interface) ............................................... 26
B  Dynccallback C library API ............................................................................. 27
   B.1  Callback Object .......................................................................................... 27
   B.2  Allocation .................................................................................................... 27
C  Dynload C library API ...................................................................................... 28
   C.1  Loading code ............................................................................................... 28
   C.2  Retrieving functions ................................................................................... 28
D  Calling Conventions .......................................................................................... 29
   D.1  x86 Calling Conventions ............................................................................. 29
      D.1.1  cdecl ........................................................................................................ 29
      D.1.2  MS fastcall ............................................................................................. 31
      D.1.3  GNU fastcall .......................................................................................... 33
      D.1.4  Borland fastcall ....................................................................................... 34
      D.1.5  Watcom fastcall ...................................................................................... 35
      D.1.6  win32 stdcall ......................................................................................... 36
      D.1.7  MS thiscall .............................................................................................. 37
      D.1.8  GNU thiscall ........................................................................................... 39
      D.1.9  pascal ....................................................................................................... 40
      D.1.10 plan9call ................................................................................................. 41
   D.2  x64 Calling Convention .............................................................................. 43
      D.2.1  MS Windows ............................................................................................ 43
      D.2.2  System V (Linux / *BSD / MacOS X) .................................................... 45
   D.3  PowerPC (32bit) Calling Convention .......................................................... 47
      D.3.1  Mac OS X/Darwin ................................................................................... 47
      D.3.2  System V PPC 32-bit ............................................................................... 49
   D.4  ARM32 Calling Convention ........................................................................ 52
      D.4.1  ATPCS ARM mode ................................................................................ 52
      D.4.2  ATPCS THUMB mode .......................................................................... 54
      D.4.3  EABI (ARM and THUMB mode) ............................................................. 56
      D.4.4  ARM on Apple’s iOS (Darwin) Platform ................................................. 57
      D.4.5  Architectures ......................................................................................... 57
   D.5  MIPS Calling Convention ........................................................................... 58
      D.5.1  MIPS EABI 32-bit Calling Convention .................................................. 58
E  Literature ........................................................................................................... 60
Listings

1. C function call ................................................................. 6
2. Assembly X86 32-bit function call ........................................ 6
3. Foreign function call in C .................................................. 8
4. Dyncall C library example .................................................. 8
5. Dyncall Python bindings example ........................................ 8
6. Dyncall R bindings example ................................................ 8
1 Motivation

Interoperability between programming languages is a desirable feature in complex software systems. While functions in scripting languages and virtual machine languages can be called in a dynamic manner, statically compiled programming languages such as C, C++ and Objective-C lack this ability. The majority of systems use C function interfaces as their system-level interface. Calling these (foreign) functions from within a dynamic environment often involves the development of so called "glue code" on both sides, the use of external tools generating communication code, or integration of other middleware fulfilling that purpose. However, even inside a completely static environment, without having to bridge multiple languages, it can be very useful to call functions dynamically. Consider, for example, message systems, dynamic function call dispatch mechanisms, without even knowing about the target.

The dyncall library project provides a clean and portable C interface to dynamically issue calls to foreign code using small call kernels written in assembly. Instead of providing code for every bridged function call, which unnecessarily results in code bloat, only a modest number of instructions are used to invoke all the calls.

1.1 Static function calls in C

The C programming language and its direct derivatives are limited in the way function calls are handled. A C compiler regards a function call as a fully qualified atomic operation. In such a statically typed environment, this includes the function call’s argument arity and type, as well as the return type.

1.2 Anatomy of machine-level calls

The process of calling a function on the machine level yields a common pattern:

1. The target function’s calling convention dictates how the stack is prepared, arguments are passed, results are returned and how to clean up afterwards.

2. Function call arguments are loaded in registers and on the stack according to the calling convention that take alignment constraints into account.

3. Control flow transfer from caller to callee.

4. Process return value, if any. Some calling conventions specify that the caller is responsible for cleaning up the argument stack.
The following example depicts a C source and the corresponding assembly for the X86 32-bit processor architecture.

```c
extern void f(int x, double y, float z);
void caller()
{
    f(1, 2.0, 3.0f);
}
```

**Listing 1: C function call**

```asm
.global f  ; external symbol 'f'
caller:
push 40400000H ; 3.0f (32 bit float)
              ; 2.0 (64 bit float)
push 40000000H ; low DWORD
push 0H        ; high DWORD
push 1H        ; 1 (32 bit integer)
call f         ; call 'f'
add esp, 16    ; cleanup stack
```

**Listing 2: Assembly X86 32-bit function call**
2 Overview

The dyncall library encapsulates architecture-, OS- and compiler-specific function call semantics in a virtual

\textit{bind argument parameters from left to right and then call}

interface allowing programmers to call C functions in a completely dynamic manner. In other words, instead of calling a function directly, the dyncall library provides a mechanism to push the function parameters manually and to issue the call afterwards.

Since the idea behind this concept is similar to call dispatching mechanisms of virtual machines, the object that can be dynamically loaded with arguments, and then used to actually invoke the call, is called CallVM. It is possible to change the calling convention used by the CallVM at run-time. Due to the fact that nearly every platform comes with one or more calling conventions, the dyncall library project intends to be an open-source approach to the variety of compiler-specific binary interfaces, platform specific subtleties, and so on...

The core of the library consists of dynamic implementations of different calling conventions written in assembler. Although the library aims to be highly portable, some assembler code needs to be written for nearly every platform/compiler/OS combination. Unfortunately, there are architectures we just don’t have at home or work. If you want to see dyncall running on such a platform, feel free to send in code and patches, or even to donate hardware you don’t need anymore. Check the \textbf{supported platforms} section for an overview of the supported platforms and the different calling convention sections for details about the support.

2.1 Features

- A portable and extendable function call interface for the C programming language.
- Ports to major platforms including Windows, Mac OS X, Linux, BSD derivates, Playstation Portable and Nintendo DS.
- Add-on language bindings to Python,R,Ruby.
- High-level state machine design using C to model calling convention parameter transfer.
- One assembly \textit{hybrid} call routine per calling convention.
- Formatted call, ellipsis function API.
- Comprehensive test suite.
2.2 Showcase

Foreign function call in C

This section demonstrates how the foreign function call is issued without, and then with, the help of the *dyncall* library and scripting language bindings.

```c
double call_as_sqrt ( void * funptr , double x)
{
    return ( ( double (*)( double ) ) funptr )(x);
}
```

Listing 3: Foreign function call in C

Dyncall C library example

The same operation can be broken down into atomic pieces (specify calling convention, binding arguments, invoking the call) using the *dyncall* library.

```c
#include <dyncall.h>
double call_as_sqrt ( void * funptr , double x)
{
    double r;
    DCCallVM* vm = dcNewCallVM(4096);
    dcMode (vm , DC_CALL_C_DEFAULT );
    dcReset (vm );
    dcArgDouble (vm , x);
    r = dcCallDouble (vm , funptr );
    dcFree (vm );
    return r;
}
```

Listing 4: Dyncall C library example

Python example

```python
import pydc
def call_as_sqrt ( funptr ,x):
    return pydc . call ( funptr , "d)d", x)
```

Listing 5: Dyncall Python bindings example

R example

```r
library(rdyncall)
call.as.sqrt <- function(funptr,x)
    .dyncall(funptr,"d)d", x)
```

Listing 6: Dyncall R bindings example
2.3 Supported platforms/architectures

The feature matrix below gives a brief overview of the currently supported platforms. Different colors are used, where a green cell indicates a supported platform, yellow a platform that might work (but is untested) and red a platform that is currently unsupported. Gray cells are combinations that don’t exist at the time of writing, or that are not taken into account.

Please note that a green cell doesn’t imply that all existing calling conventions/features/build tools are supported for that platform (but the most important). For details about the support consult the appendix.

Table 1: Supported platforms
3 Building the library

The library has been built and used successfully on several platform/architecture configurations and build systems. Please see notes on specific platforms to check if the target architecture is currently supported.

3.1 Requirements

The following tools are supported directly to build the dyncall library. However, as the number of source files to be compiled for a given platform is small, it shouldn’t be difficult to build it manually with another toolchain.

- C compiler to build the dyncall library (GCC or Microsoft C/C++ compiler)
- C++ compiler to build the optional test cases (GCC or Microsoft C/C++ compiler)
- BSD make, GNU make, or Microsoft nmake as automated build tools
- Python (optional - for generation of some test cases)
- Lua (optional - for generation of some test cases)
- CMake (optional support)

3.2 Supported/tested platforms and build systems

Although it is possible to build the dyncall library on more platforms than the ones outlined here, this section doesn’t list operating systems or architectures the authors didn’t test. However, untested platforms using the same build tools (e.g. the BSD family of operating systems using similar flavors of the BSD make utility along with GCC, etc.) should work without modification. If you have problems building the dyncall library on one of the platforms mentioned below, or if you successfully built it on a yet unlisted one, please let us know.
### x86

<table>
<thead>
<tr>
<th>Platform</th>
<th>Make Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows</td>
<td>nmake, GNU make (via Cygwin, MSYS, unixutils)</td>
</tr>
<tr>
<td>Darwin</td>
<td>GNU make, BSD make</td>
</tr>
<tr>
<td>Linux</td>
<td>GNU make</td>
</tr>
<tr>
<td>Solaris</td>
<td>GNU make (Sun’s make tool isn’t supported)</td>
</tr>
<tr>
<td>FreeBSD</td>
<td>BSD make</td>
</tr>
<tr>
<td>NetBSD</td>
<td>BSD make</td>
</tr>
<tr>
<td>OpenBSD</td>
<td>BSD make</td>
</tr>
<tr>
<td>DragonFlyBSD</td>
<td>BSD make</td>
</tr>
<tr>
<td>Plan9</td>
<td>mk</td>
</tr>
<tr>
<td>Haiku/BeOS</td>
<td>GNU make</td>
</tr>
</tbody>
</table>

### x64

<table>
<thead>
<tr>
<th>Platform</th>
<th>Make Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Windows</td>
<td>nmake</td>
</tr>
<tr>
<td>Darwin</td>
<td>GNU make, BSD make</td>
</tr>
<tr>
<td>Linux</td>
<td>GNU make</td>
</tr>
<tr>
<td>Solaris</td>
<td>GNU make (Sun’s make tool isn’t supported)</td>
</tr>
<tr>
<td>FreeBSD</td>
<td>BSD make</td>
</tr>
<tr>
<td>NetBSD</td>
<td>BSD make</td>
</tr>
<tr>
<td>OpenBSD</td>
<td>BSD make</td>
</tr>
</tbody>
</table>

### PowerPC (32bit)

<table>
<thead>
<tr>
<th>Platform</th>
<th>Make Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Darwin</td>
<td>GNU make, BSD make</td>
</tr>
<tr>
<td>Linux</td>
<td>GNU make</td>
</tr>
<tr>
<td>NetBSD</td>
<td>BSD make</td>
</tr>
</tbody>
</table>

### ARM32

<table>
<thead>
<tr>
<th>Platform</th>
<th>Make Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>NetBSD</td>
<td>BSD make</td>
</tr>
<tr>
<td>Linux</td>
<td>GNU make</td>
</tr>
<tr>
<td>Nintendo DS</td>
<td>nmake (and devkitPro tools)</td>
</tr>
<tr>
<td>iOS/iPhone</td>
<td>GNU make (and iPhone SDK on Mac OS X)</td>
</tr>
</tbody>
</table>

### MIPS32

<table>
<thead>
<tr>
<th>Platform</th>
<th>Make Tools</th>
</tr>
</thead>
<tbody>
<tr>
<td>Playstation Portable</td>
<td>GNU make (and psptoolchain tools)</td>
</tr>
</tbody>
</table>
3.3 Build instructions

1. Configure the source

*nix flavour

./configure [--option ...]

windows flavour

\.\configure [/option ...]

plan9 flavour

/>.configure.rc [--option ...]

Available options (not on Plan9, though):

help                      display help
prefix=path               specify installation prefix (Unix shell)
prefix path               specify installation prefix (Windows batch)
prefix-bd=path            specify build directory prefix (Unix shell)
prefix bd path            specify build directory prefix (Windows batch)
target-x86                build for x86 architecture
target-x64                build for x64 architecture
target-ppc32              build for ppc 32-bit architecture (not on windows batch)
target-psp                 cross-compile build for Playstation Portable (homebrew SDK)
target-arm-arm            build for ARM architecture (using ARM mode)
target-arm-thumb          build for ARM architecture (using THUMB mode)
target-nds-arm            cross-compile build for Nintendo DS (using ARM mode)
target-nds-thumb          cross-compile build for Nintendo DS (using THUMB mode)
target-cygwin             build for Cygwin platform
target-windows            build for Windows platform (MinGW, etc.)
target-macosx              build for Mac OS X platform
target-iphoneos           build for iPhone OS platform
target-universal          build universal binaries (for Mac OS X/Darwin)
tool-gcc                  use GNU Compiler Collection tool-chain
tool-msvc                 use Microsoft Visual C++
tool-pspsdk               use PSP SDK tool-chain
asm-as                    use the GNU Assembler
asm-nasm                  use NASM Assembler
asm-ml                    use Microsoft Macro Assembler
with-iphonesdk=version    use iPhone SDK version
config-release            build release version (default)
config-debug              build debug version

2. Build the static libraries dyncall, dynload and dyncallback

make                      # when using {GNU,BSD} Make
bsdmake                    # when using BSD Make on Darwin
make -f BSDmakefile       # when using BSD Make on NetBSD
nmake /f Nmakefile        # when using NMake on Windows
mk                         # when mk on Plan9
3. Install libraries and includes (supported for GNU make based builds, only)

```bash
make install
```

4. Optionally, build the test suites

```bash
make test  # when using {GNU, BSD} Make
bsdmake test  # when using BSD Make on Darwin
make -f BSDmakefile test  # when using BSD Make on NetBSD
nmake /f Nmakefile test  # when using NMake on Windows
mk test  # when using mk on Plan9
```

5. Optionally, build the manual (LaTeX required)

```bash
make doc  # when using {GNU, BSD} Make
bsdmake doc  # when using BSD Make on Darwin
make -f BSDmakefile doc  # when using BSD Make on NetBSD
nmake /f Nmakefile doc  # when using NMake on Windows
mk doc  # when using mk on Plan9
```

3.4 Build-tool specific notes

Problem: On windows using mingw and msys/unixutils 'Make', the make uses 'cc' for C compilation, which does not exist in mingw.

Solution: Set and export the 'CC' environment variable explicitly to 'gcc'.

3.5 Build with CMake

```bash
cmake -DCMAKE_INSTALL_PREFIX=<location>
make -f Makefile
```

The `-f Makefile` is required. Otherwise buildsys/gmake GNUmakefile's would be used per default.
4 Bindings to programming languages

Through binding of the dyncall library into a scripting environment, the scripting language can gain system programming status to a certain degree.
The dyncall library provides bindings to Java[1], Lua[2], Python[3], R[4] and Ruby[5].
However, please note that some of these bindings are work-in-progress and not automatically tested, meaning it might require some additional work to make them work.

4.1 Common Architecture

The binding interface of the dyncall library to various scripting languages share a common set of functionality to invoke a function call.

4.1.1 Dynamic loading of code

The helper library dynload which accompanies the dyncall library provides an abstract interface to operating-system specific mechanisms for loading a code module.

4.1.2 Functions

All bindings are based on a common interface convention providing a common set of the following 4 functions:

load loading a module of compiled code
free unloading a module of compiled code
find finding function pointer by symbolic names
call invoking a function call
4.1.3 Signatures

A signature is a character string that represents a function’s arguments and return value types. It is used in the scripting language bindings invoke functions to perform automatic type-conversion of the languages’ types to the low-level C/C++ data types. The high-level C interface functions dcCallF() and dcCallFV() also make use of the dyncall signature string.

The format of a dyncall signature string is as depicted below:

**dyncall signature string format**

\[<\text{input parameter type signature character}>^* \text{‘} ) <\text{return type signature character}>\]

The \(<\text{input parameter type signature character}>^* \text{‘} )\) sequence left to the ‘)’ is in left-to-right order of the corresponding C function parameter type list.
The special \(<\text{return type signature character}>\text{‘} v\)’ specifies that the function does not return a value and corresponds to void functions in C.

<table>
<thead>
<tr>
<th>Signature character</th>
<th>C/C++ data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>'B'</td>
<td>Bool, bool</td>
</tr>
<tr>
<td>'c'</td>
<td>char</td>
</tr>
<tr>
<td>'C'</td>
<td>unsigned char</td>
</tr>
<tr>
<td>'s'</td>
<td>short</td>
</tr>
<tr>
<td>'S'</td>
<td>unsigned short</td>
</tr>
<tr>
<td>'I'</td>
<td>int</td>
</tr>
<tr>
<td>'I'</td>
<td>unsigned int</td>
</tr>
<tr>
<td>'j'</td>
<td>long</td>
</tr>
<tr>
<td>'J'</td>
<td>unsigned long</td>
</tr>
<tr>
<td>'l'</td>
<td>long long, int64_t</td>
</tr>
<tr>
<td>'L'</td>
<td>unsigned long long, uint64_t</td>
</tr>
<tr>
<td>'f'</td>
<td>float</td>
</tr>
<tr>
<td>'d'</td>
<td>double</td>
</tr>
<tr>
<td>'p'</td>
<td>void*</td>
</tr>
<tr>
<td>'Z'</td>
<td>const char* (pointing to C string)</td>
</tr>
<tr>
<td>'v'</td>
<td>void</td>
</tr>
</tbody>
</table>

Table 2: Type signature encoding for function call data types

While the size and encoding scheme (integer, float or double) is an important property for the dyncall library to establish the function in a correct way, the distinction between signed and unsigned integer data types (char, short, int, long, long long) in C is not of importance for the library itself, as these types share the same machine storage semantics in regard to register usage, size and alignment.

On a higher level, such as in the binding of a scripting environment, it is vital to have a correct conversion between different storage schemas among scripting languages. Therefore we define also the unsigned/signed variants for a later use in the language binding part.
Examples of C function prototypes

<table>
<thead>
<tr>
<th>C function prototype</th>
<th>dyncall signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>void f1();</td>
<td>&quot;)v&quot;</td>
</tr>
<tr>
<td>int f2(int, int);</td>
<td>&quot;)ii&quot;</td>
</tr>
<tr>
<td>long long f3(void*);</td>
<td>&quot;)pL&quot;</td>
</tr>
<tr>
<td>double f4(int, bool, char, double, const char*);</td>
<td>&quot;)BcdZ&quot;</td>
</tr>
</tbody>
</table>

Table 3: Type signature examples of C function prototypes

4.2 Python language bindings

The python module pydc implements the Python language bindings, namely load, find, free, call.

<table>
<thead>
<tr>
<th>Signature character</th>
<th>accepted Python data types</th>
</tr>
</thead>
<tbody>
<tr>
<td>'B'</td>
<td>bool</td>
</tr>
<tr>
<td>'c'</td>
<td>if string, take first item</td>
</tr>
<tr>
<td>'s'</td>
<td>int, check in range</td>
</tr>
<tr>
<td>'i'</td>
<td>int</td>
</tr>
<tr>
<td>'j'</td>
<td>int</td>
</tr>
<tr>
<td>'l'</td>
<td>long, casted to long long</td>
</tr>
<tr>
<td>'f'</td>
<td>float</td>
</tr>
<tr>
<td>'d'</td>
<td>double</td>
</tr>
<tr>
<td>'p'</td>
<td>string or long casted to void*</td>
</tr>
<tr>
<td>'v'</td>
<td>no return type</td>
</tr>
</tbody>
</table>

Table 4: Type signature encoding for Python bindings

4.3 R language bindings

The R package rdyncall implements the R language bindings providing the function .dyncall().

Some notes on the R Binding:

- Unsigned 32-bit integers are represented as signed integers in R.
- 64-bit integer types do not exist in R, therefore we use double floats to represent 64-bit integers (using only the 52-bit mantissa part).
### Signature character accepted R data types

<table>
<thead>
<tr>
<th>Signature character</th>
<th>Accepted R data types</th>
</tr>
</thead>
<tbody>
<tr>
<td>'B'</td>
<td>coerced to logical vector, first item</td>
</tr>
<tr>
<td>'c'</td>
<td>coerced to integer vector, first item truncated char</td>
</tr>
<tr>
<td>'C'</td>
<td>coerced to integer vector, first item truncated to unsigned char</td>
</tr>
<tr>
<td>'s'</td>
<td>coerced to integer vector, first item truncated to short</td>
</tr>
<tr>
<td>'S'</td>
<td>coerced to integer vector, first item truncated to short</td>
</tr>
<tr>
<td>'i'</td>
<td>coerced to integer vector, first item</td>
</tr>
<tr>
<td>'I'</td>
<td>coerced to integer vector, first item truncated to unsigned int</td>
</tr>
<tr>
<td>'j'</td>
<td>coerced to integer vector, first item</td>
</tr>
<tr>
<td>'J'</td>
<td>coerced to integer vector, first item truncated to unsigned long</td>
</tr>
<tr>
<td>'l'</td>
<td>coerced to numeric, first item casted to long long</td>
</tr>
<tr>
<td>'L'</td>
<td>coerced to numeric, first item casted to unsigned long</td>
</tr>
<tr>
<td>'f'</td>
<td>coerced to numeric, first item casted to float</td>
</tr>
<tr>
<td>'d'</td>
<td>coerced to numeric, first item casted to float</td>
</tr>
<tr>
<td>'p'</td>
<td>external pointer or coerced to string vector, first item</td>
</tr>
<tr>
<td>'Z'</td>
<td>coerced to string vector, first item</td>
</tr>
<tr>
<td>'v'</td>
<td>no return type</td>
</tr>
</tbody>
</table>

Table 5: Type signature encoding for R bindings

### 4.4 Ruby language bindings

The Ruby gem rbdc implements the Ruby language bindings.

<table>
<thead>
<tr>
<th>Signature character</th>
<th>Accepted Ruby data types</th>
</tr>
</thead>
<tbody>
<tr>
<td>'B', 'C'</td>
<td>TrueClass, FalseClass, NilClass, Fixnum casted to bool</td>
</tr>
<tr>
<td>'s', 'S'</td>
<td>Fixnum cast to (unsigned) char</td>
</tr>
<tr>
<td>'i', 'I'</td>
<td>Fixnum cast to (unsigned) int</td>
</tr>
<tr>
<td>'j', 'J'</td>
<td>Fixnum cast to (unsigned) long</td>
</tr>
<tr>
<td>'l', 'L'</td>
<td>Fixnum cast to (unsigned) long long</td>
</tr>
<tr>
<td>'f'</td>
<td>Float cast to float</td>
</tr>
<tr>
<td>'d'</td>
<td>Float cast to double</td>
</tr>
<tr>
<td>'p', 'Z'</td>
<td>String cast to void*</td>
</tr>
<tr>
<td>'v'</td>
<td>no return type</td>
</tr>
</tbody>
</table>

Table 6: Type signature encoding for Ruby bindings
5 Library Design

5.1 Design considerations

The dyncall library encapsulates function call invocation semantics that can depend on the compiler, operating system and architecture. The core library is driven by a function call invocation engine, namely the CallVM, that encapsulates a call stack to foreign functions and manages the following three phases that constitute a dyncall function call:

1. Specify the calling convention. Some run-time platforms, such as Microsoft Windows on a 32-bit X86 architecture, even support multiple calling conventions.

2. Specify the function call arguments in a specific order. The interface design dictates a left to right order for C and C++ function calls in which the arguments are bounded.

3. Specify the target function address, expected return value and invoke the function call.

The calling convention mode entirely depends on the way the foreign function has been compiled and specifies the low-level details on how a function actually expects input parameters (in memory, in registers or both) and how to return results.
6 Developers

6.1 Project root

configure -- configuration tool (unix-shell)
configure.bat -- configuration tool (windows batch)
CMakeLists.txt -- top-level project information for CMake
ConfigVars -- configuration tool output
BSDmakefile -- BSD makefile
GNUmakefile -- GNU makefile
Nmakefile -- MS nmake makefile
LICENSE -- license information
README.txt -- general information
buildsys/ -- build systems ({BSD,GNU,N}make) and extras
doc/ -- manual
dyncall/ -- dyncall library source code
dyncallback/ -- dyncallback library source code
dynload/ -- dynload library source code
test/ -- test suites

6.2 Test suites

plain Simple, identity, unary function calls for all supported return types and calling conventions.

plain_c++ Similar to plain, but for C++ thiscalls (GNU and MS calling convention).
suite All combinations of parameter types and counts are tested on void function calls. A script written in Python (mkcase.py) generates the tests up to an upper MAXARG limit.
suite_floats Based on suite. Test double/float variants with up to 10 arguments.
suite_x86win32std All combinations of parameter types and counts are tested on _stdcall void function calls. A script written in Python (mkcase.py) generates the tests up to an upper MAXARG limit. This is a x86/Windows only test.
suite_x86win32fastcall All combinations of parameter types and counts are tested on _fastcall (MS or GNU, depending on the build tool) void function calls. A script written in Python (mkcase.py) generates the tests up to an upper MAXARG limit. This is a x86/Windows only test.
ellipsis All combinations of parameter types and counts are tested on void ellipsis function calls. A script written in Python (mkcase.py) generates the tests up to an upper MAXARG limit.
suite2 Designed mass test suite for void function calls. Tests individual void functions with a varying count of arguments and type.
suite2_win32stdcall Designed mass test suite for _stdcall void function calls. Tests individual void functions with a varying count of arguments and type. This is a x86/Windows only test.
suite2_win32fastcall Designed mass test suite for _fastcall (MS or GNU, depending on the build tool) void function calls. Tests individual void functions with a varying count of arguments and type. This is a x86/Windows only test.
suite3 All combinations of parameter types integer, long long, float and double and counts are tested on void function calls. A script written in Python (mkcase.py) generates the tests up to an upper MAXARG limit. This is a modified version of suite.
callf Tests the formatted call dyncall C API.
malloc_wx Tests writable and executable memory allocation used by the dyncallback C API.
**thunk**  Tests *callbacks* for the *dyncallback* C API.

**callback_plain**  Simple callback function test (useful for easy debugging of new ports).

**callback_suite**  Mass test suite for callback function testing. Uses random function argument number and type. A script written in Lua generates the tests up to a given number of calls and type combinations.

**resolv_self**  Test for dynload library to resolve symbols from application image itself.
7 Epilog

7.1 Stability and security considerations

Since the dyncall library doesn't know anything about the called function itself (except its address), no parameter-type validation is done. Thus in order to avoid crashes, data corruption, etc., the user is urged to ascertain the number and types of parameters. It is strongly advised to double check the parameter types of every function to be called, and not to call unknown functions at all.

Consider a simple program that issues a call by directly passing some command line arguments to the call itself, or even worse, by indirectly choosing a library and a function to call. Such unchecked input data can be quite easily used to intentionally crash the program, or to hijack it and take control of the program flow.

To put it in a nutshell, if not used with care, programs depending on the dyncall, dyncallback and dynload libraries, can become arbitrary function call dispatchers by manipulating their input data. Successful exploits of programs like the one outlined above can be misused as very powerful tools for a wide variety of malicious attacks, ...

7.2 Embedding

The dyncall library has a very low dependency to system facilities. The library uses some heap-memory to store the Call VM and uses per default malloc() and free() calls. This behaviour can be changed by providing custom dcAllocMem() and dcFreeMem() functions. See dyncall/dyncall_alloc.h for details.

7.3 Multi-threading

The dyncall library is thread-safe and reentrant, by means that it works correctly during execution of multiple threads if, and only if there is at most a single thread pushing arguments to a CallVM (invoking the call is always thread-safe, though). However, since there’s no limitation on the number of created CallVM objects, it is advised to keep a copy for each thread.

7.4 Supported types

Currently, the dyncall library supports all of ANSI C’s integer, floating point and pointer types as function call arguments as well as return values. Additionally, C++’s bool type is supported. Due to the still rare and often incomplete support of the long double type on various platforms, the latter is currently not supported.

7.5 Roadmap

The dyncall library should be extended by a wide variety of other calling conventions and ported to other, more esoteric platforms. With its low memory footprint it surely might come in handy on embedded systems. So far dyncall supports arm32 and mips32 (eabi) embedded systems processors. Furthermore, the authors plan to write some more scripting language bindings, examples, and other projects that are based on it.

Besides dyncall and dyncallback, the dynload library needs to be extended with support for other shared library formats (e.g. AmigaOS .library or GEM [28] files).

7.6 Related libraries

Besides the dyncall library, there are other free and open projects with similar goals. The most noteworthy libraries are libffi [29], C/Invoke [30] and libffcall [31].
A  Dynccall C library API

The library provides low-level functionality to make foreign function calls from different run-time environments. The flexibility is constrained by the set of supported types.

C interface style conventions

This manual and the dynccall library's C interface "dyncall.h" uses the following C source code style.

Subject | C symbol | Details | Example
--- | --- | --- | ---
Types | DC<type name> | lower-case | DCint, DCfloat, DCLong, ... | C table
Structures | DC<structure name> | camel-case | DCCallVM | C table
Functions | dc<function name> | camel-case | dcNewCallVM, dcArgInt, ... | C table

Table 7: C interface conventions

A.1  Supported C/C++ argument and return types

<table>
<thead>
<tr>
<th>Type alias</th>
<th>C/C++ data type</th>
</tr>
</thead>
<tbody>
<tr>
<td>DBCbool</td>
<td>_Bool, bool</td>
</tr>
<tr>
<td>DBCchar</td>
<td>char</td>
</tr>
<tr>
<td>DBCshort</td>
<td>short</td>
</tr>
<tr>
<td>DBCint</td>
<td>int</td>
</tr>
<tr>
<td>DBClong</td>
<td>long</td>
</tr>
<tr>
<td>DBClonglong</td>
<td>long long</td>
</tr>
<tr>
<td>DBCfloat</td>
<td>float</td>
</tr>
<tr>
<td>DBCdouble</td>
<td>double</td>
</tr>
<tr>
<td>DBCpointer</td>
<td>void*</td>
</tr>
<tr>
<td>DBCvoid</td>
<td>void</td>
</tr>
</tbody>
</table>

Table 8: Supported C/C++ argument and return types
A.2 Call Virtual Machine - CallVM

This CallVM is the main entry to the functionality of the library.

Types

```c
typedef void DCCallVM; /* abstract handle */
```

Details

The CallVM is a state machine that manages all aspects of a function call from configuration, argument passing up the actual function call on the processor.

A.3 Allocation

Functions

```c
DCCallVM* dcNewCallVM (DCsize size);
void dcFree (DCCallVM* vm);
```

dcNewCallVM creates a new CallVM object, where size specifies the size of the internal stack that will be allocated and used to bind the arguments to. Use dcFree to destroy the CallVM object.

A.4 Error Reporting

Errors

<table>
<thead>
<tr>
<th>Constant</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC_ERROR_NONE</td>
<td>No error occurred.</td>
</tr>
<tr>
<td>DC_ERROR_UNSUPPORTED_MODE</td>
<td>Unsupported mode, caused by dcMode()</td>
</tr>
</tbody>
</table>

Table 9: CallVM calling convention modes

Function

```c
DCint dcGetError (DCCallVM* vm);
```

Returns either one of the error codes, outlined, above.
A.5 Configuration

Function

```c
void dcMode (DCCallVM* vm, DCint mode);
```

Sets the calling convention to use. Note that some mode/platform combination don’t make any sense (e.g. using a PowerPC calling convention on a MIPS platform).

Modes

<table>
<thead>
<tr>
<th>Constant</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DC_CALL_C_DEFAULT</td>
<td>C default function call for current platform</td>
</tr>
<tr>
<td>DC_CALL_C_ELLIPSIS</td>
<td>C ellipsis function call (variable arguments)</td>
</tr>
<tr>
<td>DC_CALL_C_X86_CDECL</td>
<td>C x86 platforms standard call</td>
</tr>
<tr>
<td>DC_CALL_C_X86_WIN32_STD</td>
<td>C x86 Windows standard call</td>
</tr>
<tr>
<td>DC_CALL_C_X86_WIN32_FAST_MS</td>
<td>C x86 Windows Microsoft fast call</td>
</tr>
<tr>
<td>DC_CALL_C_X86_WIN32_FAST_GNU</td>
<td>C x86 Windows GCC fast call</td>
</tr>
<tr>
<td>DC_CALL_C_X86_WIN32_THIS_MS</td>
<td>C x86 Windows Microsoft this call</td>
</tr>
<tr>
<td>DC_CALL_C_X86_WIN32_THIS_GNU</td>
<td>C x86 Windows GCC this call</td>
</tr>
<tr>
<td>DC_CALL_C_X86_PLAN9</td>
<td>C x86 Plan9 call</td>
</tr>
<tr>
<td>DC_CALL_C_X64_WIN64</td>
<td>C x64 Windows standard call</td>
</tr>
<tr>
<td>DC_CALL_C_X64_SYSV</td>
<td>C x64 System V standard call</td>
</tr>
<tr>
<td>DC_CALL_C_PPC32_DARWIN</td>
<td>C ppc32 Mac OS X standard call</td>
</tr>
<tr>
<td>DC_CALL_C_PPC32_SYSV</td>
<td>C ppc32 System V standard call</td>
</tr>
<tr>
<td>DC_CALL_C_ARM_ARM</td>
<td>C arm call (arm mode)</td>
</tr>
<tr>
<td>DC_CALL_C_ARM_THUMB</td>
<td>C arm call (thumb mode)</td>
</tr>
<tr>
<td>DC_CALL_C_ARM_ARM_EABI</td>
<td>C arm eabi call (arm mode)</td>
</tr>
<tr>
<td>DC_CALL_C_ARM_THUMB_EABI</td>
<td>C arm eabi call (thumb mode)</td>
</tr>
<tr>
<td>DC_CALL_C_MIPS32_EABI</td>
<td>C mips32 eabi call</td>
</tr>
<tr>
<td>DC_CALL_C_MIPS32_PSPSDK</td>
<td>C mips32 default PSP Homebrew SDK call (uses eabi)</td>
</tr>
<tr>
<td>DC_CALL_C_MIPS32_032</td>
<td>C mips32 o32 call</td>
</tr>
<tr>
<td>DC_CALL_C_MIPS64_N64</td>
<td>C mips64 n64 call</td>
</tr>
<tr>
<td>DC_CALL_C_MIPS64_N32</td>
<td>C mips64 n32 call</td>
</tr>
</tbody>
</table>

Table 10: CallVM calling convention modes

Details

DC_CALL_C_DEFAULT is the default standard C call on the target platform. It uses the standard C calling convention. DC_CALL_C_ELLIPSIS is used for C ellipsis calls which allow to build up a variable argument list. On many platforms, there is only one C calling convention. The X86 platform provides a rich family of different calling conventions.
A.6 Machine state reset

```c
void dcReset(DCCallVM* vm);
```

Resets the internal stack of arguments and prepares it for the selected mode. This function should be called after setting the call mode (using dcMode), but prior to binding arguments to the CallVM. Use it when reusing a CallVM, too, as arguments don’t get flushed automatically after a function call invocation.

A.7 Argument binding

**Functions**

- `void dcArgBool (DCCallVM* vm, DCbool arg);`
- `void dcArgChar (DCCallVM* vm, DCchar arg);`
- `void dcArgShort (DCCallVM* vm, DCshort arg);`
- `void dcArgInt (DCCallVM* vm, DCint arg);`
- `void dcArgLong (DCCallVM* vm, DClong arg);`
- `void dcArgLongLong (DCCallVM* vm, DClonglong arg);`
- `void dcArgFloat (DCCallVM* vm, DCfloat arg);`
- `void dcArgDouble (DCCallVM* vm, DCdouble arg);`
- `void dcArgPointer (DCCallVM* vm, DCpointer arg);`

**Details**

Used to bind arguments of the named types to the CallVM object. Arguments should be bound in left-to-right order regarding the C function prototype.

A.8 Call invocation

**Functions**

- `DCvoid dcCallVoid (DCCallVM* vm, DCpointer funcptr);`
- `DCbool dcCallBool (DCCallVM* vm, DCpointer funcptr);`
- `DCchar dcCallChar (DCCallVM* vm, DCpointer funcptr);`
- `DCshort dcCallShort (DCCallVM* vm, DCpointer funcptr);`
- `DCint dcCallInt (DCCallVM* vm, DCpointer funcptr);`
- `DClong dcCallLong (DCCallVM* vm, DCpointer funcptr);`
- `DClonglong dcCallLongLong (DCCallVM* vm, DCpointer funcptr);`
- `DCfloat dcCallFloat (DCCallVM* vm, DCpointer funcptr);`
- `DCdouble dcCallDouble (DCCallVM* vm, DCpointer funcptr);`
- `DCpointer dcCallPointer (DCCallVM* vm, DCpointer funcptr);`

**Details**

After the invocation of the foreign function call, the argument values are still bound and a second call using the same arguments can be issued. If you need to clear the argument bindings, you have to reset the CallVM.
A.9 Formatted calls (ANSI C ellipsis interface)

Functions

```c
void dcCallF (DCCallVM* vm, DCValue* result, DCpointer funcptr,
              const DCsigchar* signature, ...);
void dcVCallF (DCCallVM* vm, DCValue* result, DCpointer funcptr,
               const DCsigchar* signature, va_list args);
```

Details

These functions can be used to operate dyncall via a printf-style functional interface, using a signature string encoding the argument types and return type. The return value will be stored in result. For more information about the signature format, refer to [2]
B  DynCallback C library API

This library extends dyncall with function callback support, allowing the user to dynamically create a callback object that can be called directly, or passed to functions expecting a function-pointer as argument. Once again, the flexibility is constrained by the set of supported types.

For style conventions and supported types, see dyncall API section. In order to use dyncallback, include "dyncall_callback.h".

B.1 Callback Object

The Callback Object is the core component to this library.

Types

```c
typedef struct DCCallback DCCallback;
```

Details

The Callback Object is an object that mimics a fully typed function call to another function (a generic callback handler, in this case). The object itself can be created dynamically.

B.2 Allocation

Functions

```c
DCCallback* dcbNewCallback(const char* signature, DCCallbackHandler* funcptr, void* userdata);
void dcbFreeCallback(DCCallback* pcb);
```

dcbNewCallback creates and initializes a new Callback object, where signature is the needed function signature (format is the one outlined in the language bindings-section of this manual) of the function to mimic, funcptr is a pointer to a callback handler, and userdata a pointer to custom data that might be useful in the handler. Use dcbFreeCallback to destroy the Callback object.
C  Dynload C library API

The *dynload* library encapsulates dynamic loading mechanisms and gives access to functions in foreign dynamic libraries and code modules.

C.1  Loading code

```c
void* dlLoadLibrary(const char* libpath);
void  dlFreeLibrary(void* libhandle);
```

C.2  Retrieving functions

```c
void* dlFindSymbol(void* libhandle, const char* symbol);
```
D Calling Conventions

Before we go any further...

It is of great importance to be aware that this section isn't a general purpose description of the present calling conventions. It merely explains the calling conventions for the parameter/return types supported by dyncall, not for aggregates (structures, unions and classes), SIMD data types (\_m64, \_m128, \_m128i, \_m128d), etc. We strongly advise the reader not to use this document as a general purpose calling convention reference.

D.1 x86 Calling Conventions

Overview

There are numerous different calling conventions on the x86 processor architecture, like cdecl [6], MSfastcall [8], GNUfastcall [9], Borlandfastcall [10], Watcomfastcall [11], Win32stdcall [7], MSthiscall [12], GNUthiscall [13], the pascal calling convention [14] and a cdecl-like version for Plan9 [15] (dubbed plan9call by us), etc.

dyncall support

Currently cdecl, stdcall,fastcall (MS and GNU), thiscall (MS and GNU) and plan9call are supported.

D.1.1 cdecl

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>permanent</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch</td>
</tr>
<tr>
<td>edx</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>esi</td>
<td>permanent</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 11: Register usage on x86 cdecl calling convention
Parameter passing

- stack parameter order: right-to-left
- caller cleans up the stack
- all arguments are pushed onto the stack

Return values

- return values of pointer or integral type (<= 32 bits) are returned via the eax register
- integers > 32 bits are returned via the eax and edx registers
- floating point types are returned via the st0 register

Stack layout

Stack directly after function prolog:

Figure 1: Stack layout on x86 cdecl calling convention
### D.1.2 MS fastcall

**Registers and register usage**

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>permanent</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch, parameter 0</td>
</tr>
<tr>
<td>edx</td>
<td>scratch, parameter 1, return value</td>
</tr>
<tr>
<td>esi</td>
<td>permanent</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 12: Register usage on x86 fastcall (MS) calling convention
Parameter passing

- stack parameter order: right-to-left
- called function cleans up the stack
- first two integers/pointers (< 32bit) are passed via ecx and edx (even if preceded by other arguments)
- integer types 64 bits in size @@@ ? first in edx:eax ?
- if first argument is a 64 bit integer, it is passed via ecx and edx
- all other parameters are pushed onto the stack

Return values

- return values of pointer or integral type (< 32 bits) are returned via the eax register
- integers > 32 bits are returned via the eax and edx registers@@@verify
- floating point types are returned via the st0 register@@@ really ?

Stack layout

Stack directly after function prolog:

```
... local data
parameter area  }  stack parameters  }  caller's frame
  ... ... return address
local data
parameter area
... ...
```

Figure 2: Stack layout on x86 fastcall (MS) calling convention
D.1.3 GNU fastcall

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>permanent</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch, parameter 0</td>
</tr>
<tr>
<td>edx</td>
<td>scratch, parameter 1, return value</td>
</tr>
<tr>
<td>esi</td>
<td>permanent</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 13: Register usage on x86 fastcall (GNU) calling convention

Parameter passing

- stack parameter order: right-to-left
- called function cleans up the stack
- first two integers/pointers (\(\leq 32\text{bit}\)) are passed via ecx and edx (even if preceded by other arguments)
- if first argument is a 64 bit integer, it is pushed on the stack and the two registers are skipped
- all other parameters are pushed onto the stack

Return values

- return values of pointer or integral type (\(\leq 32\text{ bits}\)) are returned via the eax register.
- integers > 32 bits are returned via the eax and edx registers.
- floating point types are returned via the st0.
Stack layout

Stack directly after function prolog:

```
local data

parameter area

    ... return address

    ... stack parameters

caller's frame

current frame

local data
```

Figure 3: Stack layout on x86 fastcall (GNU) calling convention

D.1.4 Borland fastcall

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, parameter 0, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>permanent</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch, parameter 2</td>
</tr>
<tr>
<td>edx</td>
<td>scratch, parameter 1, return value</td>
</tr>
<tr>
<td>esi</td>
<td>permanent</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 14: Register usage on x86 fastcall (Borland) calling convention

Parameter passing

- stack parameter order: left-to-right
- called function cleans up the stack
- first three integers/pointers ($\leq 32$bit) are passed via eax, ecx and edx (even if preceded by other arguments)
- integer types 64 bits in size
- all other parameters are pushed onto the stack
Return values

- return values of pointer or integral type (≤ 32 bits) are returned via the eax register
- integers > 32 bits are returned via the eax and edx registers@@@ verify
- floating point types are returned via the st0 register@@@ really?

Stack layout

Stack directly after function prolog:

```
local data

parameter area

stack parameters

return address

local data

parameter area

return address

stack parameters

caller's frame

current frame
```

Figure 4: Stack layout on x86 fastcall (Borland) calling convention

D.1.5 Watcomfastcall

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, parameter 0, return value@@@</td>
</tr>
<tr>
<td>ebx</td>
<td>scratch when used for parameter, parameter 2</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch when used for parameter, parameter 3</td>
</tr>
<tr>
<td>edx</td>
<td>scratch when used for parameter, parameter 1, return value@@@</td>
</tr>
<tr>
<td>esi</td>
<td>scratch when used for return pointer @@??</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 15: Register usage on x86fastcall (Watcom) calling convention

Parameter passing

- stack parameter order: right-to-left
- called function cleans up the stack
• first four integers/pointers (≤ 32bit) are passed via eax, edx, ebx and ecx (even if preceded by other arguments?)

• integer types 64 bits in size ??

• all other parameters are pushed onto the stack

Return values

• return values of pointer or integral type (≤ 32 bits) are returned via the eax register??verify, I think its esi?

• integers > 32 bits are returned via the eax and edx registers?? verify

• floating point types are returned via the st0 register?? really ?

Stack layout

Stack directly after function prolog:

```
local data
|
parameter area { ... stack parameters } caller’s frame
|
local data
|
parameter area
```

Figure 5: Stack layout on x86 fastcall (Watcom) calling convention

D.1.6 win32 stdcall

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>permanent</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch</td>
</tr>
<tr>
<td>edx</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>esi</td>
<td>permanent</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 16: Register usage on x86 stdcall calling convention
Parameter passing

- Stack parameter order: right-to-left
- Called function cleans up the stack
- All parameters are pushed onto the stack
- Stack is usually 4 byte aligned (GCC >= 3.x seems to use a 16byte alignment
- Function name is decorated by prepending an underscore character and appending a 'O' character and the number of bytes of stack space required

Return values

- return values of pointer or integral type (<= 32 bits) are returned via the eax register
- integers > 32 bits are returned via the eax and edx registers
- floating point types are returned via the st0 register

Stack layout

Stack directly after function prolog:

```
... local data ...
parameter area  }
               }
stack parameters caller's frame
           return address
... local data ...
parameter area  }
               }
current frame
       ...
```

Figure 6: Stack layout on x86 stdcall calling convention

D.1.7 MS thiscall

Registers and register usage
<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>permanent</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch, parameter 0</td>
</tr>
<tr>
<td>edx</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>esi</td>
<td>permanent</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 17: Register usage on x86 thiscall (MS) calling convention

Parameter passing

- stack parameter order: right-to-left
- called function cleans up the stack
- first parameter (this pointer) is passed via ecx
- all other parameters are pushed onto the stack
- Function name is decorated by prepending a ’@’ character and appending a ’@’ character and the number of bytes (decimal) of stack space required

Return values

- return values of pointer or integral type (<= 32 bits) are returned via the eax register
- integers > 32 bits are returned via the eax and edx registers
- floating point types are returned via the st0 register really?

Stack layout

Stack directly after function prolog:

![Stack layout diagram](image-url)

Figure 7: Stack layout on x86 thiscall (MS) calling convention
D.1.8 GNU thiscall

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>permanent</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch</td>
</tr>
<tr>
<td>edx</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>esi</td>
<td>permanent</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 18: Register usage on x86 thiscall (GNU) calling convention

Parameter passing

- stack parameter order: right-to-left
- caller cleans up the stack
- all parameters are pushed onto the stack

Return values

- return values of pointer or integral type (<= 32 bits) are returned via the eax register
- integers > 32 bits are returned via the eax and edx registers
- floating point types are returned via the st0 register really?

Stack layout

Stack directly after function prolog:

```
local data
...
parameter area { ... } stack parameters { caller's frame
  ...
  return address
  ...
local data
parameter area
  ...
```  

Figure 8: Stack layout on x86 thiscall (GNU) calling convention
D.1.9 pascal

The best known uses of the pascal calling convention are the 16 bit OS/2 APIs, Microsoft Windows 3.x and Borland Delphi 1.x.

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>permanent</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch</td>
</tr>
<tr>
<td>edx</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>esi</td>
<td>permanent</td>
</tr>
<tr>
<td>edi</td>
<td>permanent</td>
</tr>
<tr>
<td>ebp</td>
<td>permanent</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0</td>
<td>scratch, floating point return value</td>
</tr>
<tr>
<td>st1-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 19: Register usage on x86 pascal calling convention

Parameter passing

- stack parameter order: left-to-right
- called function cleans up the stack
- all parameters are pushed onto the stack

Return values

- return values of pointer or integral type (\(\leq 32\) bits) are returned via the eax register
- integers > 32 bits are returned via the eax and edx registers
- floating point types are returned via the st0 register

Stack layout

Stack directly after function prolog.
D.1.10 plan9call

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>eax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>ebx</td>
<td>scratch</td>
</tr>
<tr>
<td>ecx</td>
<td>scratch</td>
</tr>
<tr>
<td>edx</td>
<td>scratch</td>
</tr>
<tr>
<td>esi</td>
<td>scratch</td>
</tr>
<tr>
<td>edi</td>
<td>scratch</td>
</tr>
<tr>
<td>ebp</td>
<td>scratch</td>
</tr>
<tr>
<td>esp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>st0-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 20: Register usage on x86 plan9call calling convention

Parameter passing

- stack parameter order: right-to-left
- caller cleans up the stack
- all parameters are pushed onto the stack
Return values

- return values of pointer or integral type (\(\leq 32\) bits) are returned via the eax register
- integers \(> 32\) bits or structures are returned by the caller allocating the space and passing a pointer to the callee as a new, implicit first parameter (this means, on the stack)
- floating point types are returned via the st0 register

Stack layout

Stack directly after function prolog:

![Stack layout diagram]

Figure 10: Stack layout on x86 plan9call calling convention
D.2 x64 Calling Convention

Overview

The x64 (64bit) architecture designed by AMD is based on Intel’s x86 (32bit) architecture, supporting it natively. It is sometimes referred to as x86-64, AMD64, or, cloned by Intel, EM64T or Intel64. On this processor, a word is defined to be 16 bits in size, a dword 32 bits and a qword 64 bits. Note that this is due to historical reasons (terminology didn’t change with the introduction of 32 and 64 bit processors).

The x64 calling convention for MS Windows [19] differs from the SystemV x64 calling convention [20] used by Linux/*BSD/... Note that this is not the only difference between these operating systems. The 64 bit programming model in use by 64 bit windows is LLP64, meaning that the C types int and long remain 32 bits in size, whereas long long becomes 64 bits. Under Linux/*BSD/... it’s LP64.

Compared to the x86 architecture, the 64 bit versions of the registers are called rax, rbx, etc.. Furthermore, there are eight new general purpose registers r8-r15.

dyncall support

dyncall supports the MS Windows and System V calling convention.

D.2.1 MS Windows

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>rax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>rbx</td>
<td>permanent</td>
</tr>
<tr>
<td>rcx</td>
<td>scratch, parameter 0 if integer or pointer</td>
</tr>
<tr>
<td>rdx</td>
<td>scratch, parameter 1 if integer or pointer</td>
</tr>
<tr>
<td>rdi</td>
<td>permanent</td>
</tr>
<tr>
<td>rsi</td>
<td>permanent</td>
</tr>
<tr>
<td>rbp</td>
<td>permanent, may be used as frame pointer</td>
</tr>
<tr>
<td>rsp</td>
<td>stack pointer</td>
</tr>
<tr>
<td>r8-r9</td>
<td>scratch, parameter 2 and 3 if integer or pointer</td>
</tr>
<tr>
<td>r10-r11</td>
<td>scratch, permanent if required by caller (used for syscall/sysret)</td>
</tr>
<tr>
<td>r12-r15</td>
<td>permanent</td>
</tr>
<tr>
<td>xmm0</td>
<td>scratch, floating point parameter 0, floating point return value</td>
</tr>
<tr>
<td>xmm1-xmm3</td>
<td>scratch, floating point parameters 1-3</td>
</tr>
<tr>
<td>xmm4-xmm5</td>
<td>scratch, permanent if required by caller</td>
</tr>
<tr>
<td>xmm6-xmm15</td>
<td>permanent</td>
</tr>
</tbody>
</table>

Table 21: Register usage on x64 MS Windows platform

Parameter passing

- stack parameter order: right-to-left
- caller cleans up the stack
• first 4 integer/pointer parameters are passed via rcx, rdx, r8, r9 (from left to right), others are pushed on stack (there is a preserve area for the first 4)
• float and double parameters are passed via xmm0-xmm3
• first 4 parameters are passed via the correct register depending on the parameter type - with mixed float and int parameters, some registers are left out (e.g. first parameter ends up in rcx or xmm0, second in rdx or xmm1, etc.)
• parameters in registers are right justified
• parameters < 64bits are not zero extended - zero the upper bits containing garbage if needed (but they are always passed as a qword)
• parameters > 64 bit are passed by reference
• if callee takes address of a parameter, first 4 parameters must be dumped (to the reserved space on the stack) - for floating point parameters, value must be stored in integer AND floating point register
• caller cleans up the stack, not the callee (like cdecl)
• stack is always 16byte aligned - since return address is 64 bits in size, stacks with an odd number of parameters are already aligned
• ellipsis calls take floating point values in int and float registers (single precision floats are promoted to double precision as defined for ellipsis calls)
• if size of parameters > 1 page of memory (usually between 4k and 64k), chkstk must be called

Return values
• return values of pointer or integral type (<= 64 bits) are returned via the rax register
• floating point types are returned via the xmm0 register
• for types > 64 bits, a secret first parameter with an address to the return value is passed

Stack layout
Stack frame is always 16-byte aligned. Stack directly after function prolog:
Figure 11: Stack layout on x64 Microsoft platform

Table 22: Register usage on x64 System V (Linux/*BSD)

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>rax</td>
<td>scratch, return value</td>
</tr>
<tr>
<td>rbx</td>
<td>permanent</td>
</tr>
<tr>
<td>rcx</td>
<td>scratch, parameter 3 if integer or pointer</td>
</tr>
<tr>
<td>rdx</td>
<td>scratch, parameter 2 if integer or pointer, return value</td>
</tr>
<tr>
<td>rdi</td>
<td>scratch, parameter 0 if integer or pointer</td>
</tr>
<tr>
<td>rsi</td>
<td>scratch, parameter 1 if integer or pointer</td>
</tr>
<tr>
<td>rbp</td>
<td>permanent, may be used as frame pointer</td>
</tr>
<tr>
<td>r8-r9</td>
<td>scratch, parameter 4 and 5 if integer or pointer</td>
</tr>
<tr>
<td>r10-r11</td>
<td>scratch</td>
</tr>
<tr>
<td>r12-r15</td>
<td>permanent</td>
</tr>
<tr>
<td>xmm0</td>
<td>scratch, floating point parameters 0, floating point return value</td>
</tr>
<tr>
<td>xmm1-xmm7</td>
<td>scratch, floating point parameters 1-7</td>
</tr>
<tr>
<td>xmm8-xmm15</td>
<td>scratch</td>
</tr>
<tr>
<td>st0-st1</td>
<td>scratch, 16 byte floating point return value</td>
</tr>
<tr>
<td>st2-st7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Parameter passing

- stack parameter order: right-to-left
- caller cleans up the stack
- first 6 integer/pointer parameters are passed via rdi, rsi, rdx, rcx, r8, r9
- first 8 floating point parameters <= 64 bits are passed via xmm0l-xmm7l
- parameters in registers are right justified
parameters that are not passed via registers are pushed onto the stack

parameters < 64bits are not zero extended - zero the upper bits containing garbage if needed (but they are always passed as a qword)

integer/pointer parameters > 64 bit are passed via 2 registers

if callee takes address of a parameter, number of used xmm registers is passed silently in al (passed number mustn’t be exact but an upper bound on the number of used xmm registers)

stack is always 16byte aligned - since return address is 64 bits in size, stacks with an odd number of parameters are already aligned

Return values

return values of pointer or integral type (<= 64 bits) are returned via the rax register

floating point types are returned via the xmm0 register

for types > 64 bits, a secret first parameter with an address to the return value is passed - the passed in address will be returned in rax

floating point values > 64 bits are returned via st0 and st1

Stack layout

Stack frame is always 16-byte aligned. Note that there is no spill area. Stack directly after function prolog:

Figure 12: Stack layout on x64 System V (Linux/*BSD)
D.3 PowerPC (32bit) Calling Convention

Overview

- Word size is 32 bits
- Big endian (MSB) and little endian (LSB) operating modes.
- Processor operates on floats in double precision floating point arithmetic (IEEE-754) values directly (single precision is converted on the fly)
- Apple Mac OS X/Darwin PPC is specified in "Mac OS X ABI Function Call Guide". It uses Big Endian (MSB).
- Linux PPC 32-bit ABI is specified in "LSB for PPC 2.1" which is based on "System V ABI". It uses Big Endian (MSB).
- PowerPC EABI is defined in the "PowerPC Embedded Application Binary Interface 32-Bit Implementation".

dynccall support

Dynccall supports PowerPC (32bit) Big Endian (MSB) on Darwin (tested on Apple Mac OS X) and System V ABI systems (Linux, NetBSD, etc.).

D.3.1 Mac OS X/Darwin

Registers and register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>gpr0</td>
<td>scratch</td>
</tr>
<tr>
<td>gpr1</td>
<td>stack pointer</td>
</tr>
<tr>
<td>gpr2</td>
<td>scratch</td>
</tr>
<tr>
<td>gpr3</td>
<td>return value, parameter 0 if integer or pointer</td>
</tr>
<tr>
<td>gpr4-gpr10</td>
<td>return value, parameter 1-7 for integer or pointer parameters</td>
</tr>
<tr>
<td>gpr11</td>
<td>permanent</td>
</tr>
<tr>
<td>gpr12</td>
<td>branch target for dynamic code generation</td>
</tr>
<tr>
<td>gpr13-31</td>
<td>permanent</td>
</tr>
<tr>
<td>fpr0</td>
<td>scratch</td>
</tr>
<tr>
<td>fpr1-fpr13</td>
<td>parameter 0-12 for floating point (always double precision)</td>
</tr>
<tr>
<td>fpr14-fpr31</td>
<td>permanent</td>
</tr>
<tr>
<td>v0-v1</td>
<td>scratch</td>
</tr>
<tr>
<td>v2-v13</td>
<td>vector parameters</td>
</tr>
<tr>
<td>v14-v19</td>
<td>scratch</td>
</tr>
<tr>
<td>v20-v31</td>
<td>permanent</td>
</tr>
<tr>
<td>lr</td>
<td>scratch, link-register</td>
</tr>
<tr>
<td>ctr</td>
<td>scratch, count-register</td>
</tr>
<tr>
<td>cr0-cr1</td>
<td>scratch</td>
</tr>
<tr>
<td>cr2-cr4</td>
<td>permanent</td>
</tr>
<tr>
<td>cr5-cr7</td>
<td>scratch</td>
</tr>
</tbody>
</table>

Table 23: Register usage on Darwin PowerPC 32-Bit
Parameter passing

- stack parameter order: right-to-left?
- caller cleans up the stack?
- the first 8 integer parameters are passed in registers gpr3-gpr10
- the first 12 floating point parameters are passed in registers fpr1-fpr13
- if a float parameter is passed via a register, gpr registers are skipped for subsequent integer parameters (based on the size of the float - 1 register for single precision and 2 for double precision floating point values)
- the caller pushes subsequent parameters onto the stack
- for every parameter passed via a register, space is reserved in the stack parameter area (in order to spill the parameters if needed - e.g. varargs)
- ellipsis calls take floating point values in int and float registers (single precision floats are promoted to double precision as defined for ellipsis calls)
- all nonvector parameters are aligned on 4-byte boundaries
- vector parameters are aligned on 16-byte boundaries
- integer parameters < 32 bit occupy high-order bytes of their 4-byte area
- composite parameters with size of 1 or 2 bytes occupy low-order bytes of their 4-byte area. INCONSISTENT with other 32-bit PPC binary interfaces. In AIX and OS 9, padding bytes always follow the data structure
- composite parameters 3 bytes or larger in size occupy high-order bytes

Return values

- return values of integer <= 32bit or pointer type use gpr3
- 64 bit integers use gpr3 and gpr4 (hiword in gpr3, loword in gpr4)
- floating point values are returned via fpr1
- structures <= 64 bits use gpr3 and gpr4
- for types > 64 bits, a secret first parameter with an address to the return value is passed
Stack layout

Stack frame is always 16-byte aligned. Stack directly after function prolog:

![Stack layout diagram]

Figure 13: Stack layout on ppc32 Darwin

D.3.2 System V PPC 32-bit

Status

- C++ this calls do not work.
- `dyncallback` is buggy.

Registers and register usage

Parameter passing

- Stack pointer (r1) is always 16-byte aligned. The EABI differs here - it is 8-byte alignment.
- 8 general-purpose registers (r3-r10) for integer and pointer types.
- 8 floating-pointer registers (f1-f8) for float (promoted to double) and double types.
- Additional arguments are passed on the stack directly after the back-chain and saved return address (8 bytes structure) on the callers stack frame.
- 64-bit integer data types are passed in general-purpose registers as a whole in two 32-bit general purpose registers (an odd and an even e.g. r3 and r4), probably skipping an even integer register or passed on the stack. They are never splitted into a register and stack part.
- Ellipse calls set CR bit 6
<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>r0</td>
<td>scratch</td>
</tr>
<tr>
<td>r1</td>
<td>stack pointer</td>
</tr>
<tr>
<td>r2</td>
<td>system-reserved</td>
</tr>
<tr>
<td>r3-r4</td>
<td>parameter passing and return value</td>
</tr>
<tr>
<td>r5-r10</td>
<td>parameter passing</td>
</tr>
<tr>
<td>r11-r12</td>
<td>scratch</td>
</tr>
<tr>
<td>r13</td>
<td>Small data area pointer register</td>
</tr>
<tr>
<td>r14-r30</td>
<td>Local variables</td>
</tr>
<tr>
<td>r31</td>
<td>Used for local variables or environment pointer</td>
</tr>
<tr>
<td>f0</td>
<td>scratch</td>
</tr>
<tr>
<td>f1</td>
<td>parameter passing and return value</td>
</tr>
<tr>
<td>f2-f8</td>
<td>parameter passing</td>
</tr>
<tr>
<td>f9-13</td>
<td>scratch</td>
</tr>
<tr>
<td>f14-f31</td>
<td>Local variables</td>
</tr>
<tr>
<td>cr0-cr7</td>
<td>Conditional register fields, each 4-bit wide (cr0-cr1 and cr5-cr7 are scratch)</td>
</tr>
<tr>
<td>lr</td>
<td>Link register (scratch)</td>
</tr>
<tr>
<td>ctr</td>
<td>Count register (scratch)</td>
</tr>
<tr>
<td>xer</td>
<td>Fixed-point exception register (scratch)</td>
</tr>
<tr>
<td>fpscr</td>
<td>Floating-point Status and Control Register</td>
</tr>
</tbody>
</table>

Table 24: Register usage on System V ABI PowerPC Processor

Return values

- 32-bit integers use register r3, 64-bit use registers r3 and r4 (hiword in r3, loword in r4).
- floating-point values are returned using register f1.
**Stack layout**

Stack frame is always 16-byte aligned. Stack directly after function prolog:

![Stack layout diagram]

Figure 14: Stack layout on System V ABI for PowerPC 32-bit calling convention
D.4 ARM32 Calling Convention

Overview

The ARM32 family of processors is based on the Advanced RISC Machines (ARM) processor architecture (32 bit RISC). The word size is 32 bits (and the programming model is LLP64). Basically, this family of microprocessors can be run in 2 major modes:

<table>
<thead>
<tr>
<th>Mode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARM</td>
<td>32bit instruction set</td>
</tr>
<tr>
<td>THUMB</td>
<td>compressed instruction set using 16bit wide instruction encoding</td>
</tr>
</tbody>
</table>

For more details, take a look at the ARM-THUMB Procedure Call Standard (ATPCS) [16], the Procedure Call Standard for the ARM Architecture (AAPCS) [17], as well as the Debian ARM EABI port wiki [18].

dyncall support

Currently, the dyncall library supports the ARM and THUMB mode of the ARM32 family (ATPCS [16] and EABI [18]), excluding manually triggered ARM-THUMB interworking calls. Although it’s quite possible that the current implementation runs on other ARM processor families as well, please note that only the ARMv4t family has been thoroughly tested at the time of writing. Please report if the code runs on other ARM families, too.

It is important to note, that dyncall supports the ARM architecture calling convention variant with floating point hardware disabled (meaning that the FPA and the VFP (scalar mode) procedure call standards are not supported). This processor family features some instruction sets accelerating DSP and multimedia application like the ARM Jazelle Technology (direct Java bytecode execution, providing acceleration for some bytecodes while calling software code for others), etc. that are not supported by the dyncall library.

D.4.1 ATPCS ARM mode

Registers and register usage

In ARM mode, the ARM32 processor has sixteen 32 bit general purpose registers, namely r0-15:

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>r0</td>
<td>parameter 0, scratch, return value</td>
</tr>
<tr>
<td>r1</td>
<td>parameter 1, scratch, return value</td>
</tr>
<tr>
<td>r2-r3</td>
<td>parameters 2 and 3, scratch</td>
</tr>
<tr>
<td>r4-r10</td>
<td>permanent</td>
</tr>
<tr>
<td>r11</td>
<td>frame pointer, permanent</td>
</tr>
<tr>
<td>r12</td>
<td>scratch</td>
</tr>
<tr>
<td>r13</td>
<td>stack pointer, permanent</td>
</tr>
<tr>
<td>r14</td>
<td>link register, permanent</td>
</tr>
<tr>
<td>r15</td>
<td>program counter (note: due to pipeline, r15 points to 2 instructions ahead)</td>
</tr>
</tbody>
</table>

Table 25: Register usage on arm32

Parameter passing
• stack parameter order: right-to-left
• caller cleans up the stack
• first four words are passed using r0-r3
• subsequent parameters are pushed onto the stack (in right to left order, such that the stack pointer points to the first of the remaining parameters)
• if the callee takes the address of one of the parameters and uses it to address other parameters (e.g. varargs) it has to copy - in its prolog - the first four words to a reserved stack area adjacent to the other parameters on the stack
• parameters \(\leq 32\) bits are passed as 32 bit words
• 64 bit parameters are passed as two 32 bit parts (even partly via the register and partly via the stack), although this doesn’t seem to be specified in the ATPCS), with the loword coming first
• structures and unions are passed by value, with the first four words of the parameters in r0-r3
• if return value is a structure, a pointer pointing to the return value’s space is passed in r0, the first parameter in r1, etc… (see return values)

Keep the stack eight-byte aligned can improve memory access performance and is required by LDRD and STRD on ARMv5TE processors which are part of the ARM32 family, so, in order to avoid problems one should always align the stack (tests have shown, that GCC does care about the alignment when using the ellipsis)

Return values

• return values \(\leq 32\) bits use r0
• 64 bit return values use r0 and r1
• if return value is a structure, the caller allocates space for the return value on the stack in its frame and passes a pointer to it in r0

Stack layout

Stack directly after function prolog:
D.4.2 ATPCS THUMB mode

Status

- The ATPCS THUMB mode is untested.
- Ellipse calls may not work.
- C++ this calls do not work.

Registers and register usage

In THUMB mode, the ARM32 processor family supports eight 32 bit general purpose registers r0-r7 and access to high order registers r8-r15:

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>r0</td>
<td>parameter 0, scratch, return value</td>
</tr>
<tr>
<td>r1</td>
<td>parameter 1, scratch, return value</td>
</tr>
<tr>
<td>r2-r3</td>
<td>parameters 2 and 3, scratch</td>
</tr>
<tr>
<td>r4-r6</td>
<td>permanent</td>
</tr>
<tr>
<td>r7</td>
<td>frame pointer, permanent</td>
</tr>
<tr>
<td>r8-r11</td>
<td>permanent</td>
</tr>
<tr>
<td>r12</td>
<td>scratch</td>
</tr>
<tr>
<td>r13</td>
<td>stack pointer, permanent</td>
</tr>
<tr>
<td>r14</td>
<td>link register, permanent</td>
</tr>
<tr>
<td>r15</td>
<td>program counter (note: due to pipeline, r15 points to 2 instructions ahead)</td>
</tr>
</tbody>
</table>

Table 26: Register usage on arm32 thumb mode

Parameter passing

- stack parameter order: right-to-left
• caller cleans up the stack
• first four words are passed using r0-r3
• subsequent parameters are pushed onto the stack (in right to left order, such that the stack pointer points to the first of the remaining parameters)
• if the callee takes the address of one of the parameters and uses it to address other parameters (e.g. varargs) it has to copy - in its prolog - the first four words to a reserved stack area adjacent to the other parameters on the stack
• parameters \( \leq 32 \) bits are passed as 32 bit words
• 64 bit parameters are passed as two 32 bit parts (even partly via the register and partly via the stack), although this doesn't seem to be specified in the ATPCS), with the loword coming first
• structures and unions are passed by value, with the first four words of the parameters in r0-r3
• if return value is a structure, a pointer pointing to the return value's space is passed in r0, the first parameter in r1, etc. (see return values)

Return values
• return values \( \leq 32 \) bits use r0
• 64 bit return values use r0 and r1
• if return value is a structure, the caller allocates space for the return value on the stack in its frame and passes a pointer to it in r0

Stack layout

Stack directly after function prolog:
D.4.3 EABI (ARM and THUMB mode)

The ARM EABI is very similar to the ABI outlined in ARM-THUMB procedure call standard (ATPCS) [16] - however, the EABI requires the stack to be 8-byte aligned at function entries, as well as 64 bit parameters. The latter are aligned on 8-byte boundaries on the stack and 2-registers for parameters passed via register. In order to achieve such an alignment, a register might have to be skipped for parameters passed via registers, or 4-bytes on the stack for parameters passed via the stack. Refer to the Debian ARM EABI port wiki for more information [18].

Status

- The EABI THUMB mode is tested and works fine (contrary to the ATPCS).
- Ellipse calls do not work.
- C++ this calls do not work.
D.4.4 ARM on Apple's iOS (Darwin) Platform

The iOS runs on ARMv6 (iOS 2.0) and ARMv7 (iOS 3.0) architectures. Typically code is compiled in Thumb mode.

Register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>R0</td>
<td>parameter 0, scratch, return value</td>
</tr>
<tr>
<td>R1</td>
<td>parameter 1, scratch, return value</td>
</tr>
<tr>
<td>R2-R3</td>
<td>parameters 2 and 3, scratch</td>
</tr>
<tr>
<td>R4-R6</td>
<td>permanent</td>
</tr>
<tr>
<td>R7</td>
<td>frame pointer, permanent</td>
</tr>
<tr>
<td>R8</td>
<td>permanent</td>
</tr>
<tr>
<td>R9</td>
<td>permanent(iOS 2.0) and scratch (since iOS 3.0)</td>
</tr>
<tr>
<td>R10-R11</td>
<td>permanent</td>
</tr>
<tr>
<td>R12</td>
<td>scratch, intra-procedure scratch register (IP) used by dynamic linker</td>
</tr>
<tr>
<td>R13</td>
<td>stack pointer, permanent</td>
</tr>
<tr>
<td>R14</td>
<td>link register, permanent</td>
</tr>
<tr>
<td>R15</td>
<td>program counter (note: due to pipeline, r15 points to 2 instructions ahead)</td>
</tr>
<tr>
<td>CPSR</td>
<td>Program status register</td>
</tr>
<tr>
<td>D0-D7</td>
<td>scratch, aliases S0-S15, on ARMv7 also as Q0-Q3. Not accessible from Thumb mode on ARMv6.</td>
</tr>
<tr>
<td>D8-D15</td>
<td>permanent, aliases S16-S31, on ARMv7 also as Q4-A7. Not accessible from Thumb mode on ARMv6.</td>
</tr>
<tr>
<td>D16-D31</td>
<td>Only available in ARMv7, aliases Q8-Q15.</td>
</tr>
<tr>
<td>FPSCR</td>
<td>VFP status register</td>
</tr>
</tbody>
</table>

Table 27: Register usage on ARM Apple iOS

The ABI is based on the AAPCS but with some important differences listed below:

- R7 instead of R11 is used as frame pointer
- R9 is scratch since iOS 3.0, was preserved before.

D.4.5 Architectures

The ARM architecture family contains several revisions with capabilities and extensions (such as thumb-interworking and more vector registers) The following table summaries important properties of the various architecture standards.

<table>
<thead>
<tr>
<th>Arch</th>
<th>Platforms</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARMv4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ARMv4T</td>
<td>ARM 7, ARM 9, Neo FreeRunner (OpenMoko)</td>
<td></td>
</tr>
<tr>
<td>ARMv5</td>
<td></td>
<td>BLX instruction available</td>
</tr>
<tr>
<td>ARMv6</td>
<td></td>
<td>No vector registers available in thumb</td>
</tr>
<tr>
<td>ARMv7</td>
<td>iPod touch, iPhone 3GS/4</td>
<td></td>
</tr>
</tbody>
</table>

Table 28: Overview of ARM Architecture, Platforms and Details
D.5 MIPS Calling Convention

Overview

The MIPS family of processors is based on the MIPS processor architecture. Multiple revisions of the MIPS Instruction sets, namely MIPS I, MIPS II, MIPS III, MIPS IV, MIPS32 and MIPS64. Today, MIPS32 and MIPS64 for 32-bit and 64-bit respectively. Several add-on extensions exist for the MIPS family:

**MIPS-3D** simple floating-point SIMD instructions dedicated to common 3D tasks.

**MDMX** (MaDMaX) more extensive integer SIMD instruction set using 64 bit floating-point registers.

**MIPS16e** adds compression to the instruction stream to make programs take up less room (allegedly a response to the THUMB instruction set of the ARM architecture).

**MIPS MT** multithreading additions to the system similar to HyperThreading.

Unfortunately, there is actually no such thing as “The MIPS Calling Convention”. Many possible conventions are used by many different environments such as 32, O64, N32, 64 and EABI.

**dyncall support**

Currently, dyncall supports the EABI calling convention which is used on the Homebrew SDK for the Playstation Portable. As documentation for this EABI is unofficial, this port is currently experimental.

D.5.1 MIPS EABI 32-bit Calling Convention

Register usage

<table>
<thead>
<tr>
<th>Name</th>
<th>Alias</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$0</td>
<td>$zero</td>
<td>Hardware zero</td>
</tr>
<tr>
<td>$1</td>
<td>$at</td>
<td>Assembler temporary</td>
</tr>
<tr>
<td>$2-$3</td>
<td>$v0-$v1</td>
<td>Integer results</td>
</tr>
<tr>
<td>$4-$11</td>
<td>$a0-$a7</td>
<td>Integer arguments</td>
</tr>
<tr>
<td>$12-$15,$24,$25</td>
<td>$t4-$t7,$8,$9</td>
<td>Integer temporaries</td>
</tr>
<tr>
<td>$16-$23</td>
<td>$s0-$s7</td>
<td>Preserved</td>
</tr>
<tr>
<td>$26-$27</td>
<td>$kt0-$kt1</td>
<td>Reserved for kernel</td>
</tr>
<tr>
<td>$28</td>
<td>$gp</td>
<td>Global pointer</td>
</tr>
<tr>
<td>$29</td>
<td>$sp</td>
<td>Stack pointer</td>
</tr>
<tr>
<td>$30</td>
<td>$s8</td>
<td>Frame pointer</td>
</tr>
<tr>
<td>$31</td>
<td>$ra</td>
<td>Return address</td>
</tr>
<tr>
<td>hi, lo</td>
<td></td>
<td>Multiply/divide special registers</td>
</tr>
<tr>
<td>$f0,$f2</td>
<td></td>
<td>Float results</td>
</tr>
<tr>
<td>$f1,$f3,$f4-$f11,$f20-$f23</td>
<td></td>
<td>Float temporaries</td>
</tr>
<tr>
<td>$f12-$f19</td>
<td></td>
<td>Float arguments</td>
</tr>
</tbody>
</table>

Table 29: Register usage on mips32 eabi calling convention

Parameter passing

- Stack parameter order: right-to-left
- Caller cleans up the stack
• Stack always aligned to 8 bytes.
• first 8 integers and floats are passed independently in registers using $a0-$a7 and $f12-$f19, respectively.
• if either integer or float registers are consumed up, the stack is used.
• 64-bit floats and integers are passed on two integer registers starting at an even register number, probably skipping one odd register.
• $a0-$a7 and $f12-$f19 are not required to be preserved.
• results are returned in $v0 (32-bit integer), $v0 and $v1 (64-bit integer/float), $f0 (32 bit float) and $f0 and $f2 (2 \times 32\text{ bit float e.g. complex}).

Stack layout
Stack directly after function prolog:

![Stack layout](image)

Figure 17: Stack layout on mips32 eabi calling convention
E Literature

References

[1] Java Programming Language
   http://www.java.com/

[2] The Programming Language Lua
   http://www.lua.org/

[3] Python Programming Language
   http://www.python.org/

   http://www.r-project.org/

[5] Ruby Programming Language
   http://www.ruby-lang.org/

[6] cdecl calling convention / Calling conventions on the x86 platform
   http://en.wikipedia.org/wiki/X86_calling_conventions#cdecl
   http://www.angelcode.com/dev/callconv/callconv.html#thiscall

[7] Windows stdcall calling convention / Microsoft calling conventions

[8] Windowsfastcall calling convention / Microsoft calling conventions

[9] GNUfastcall calling convention / Calling conventions on the x86 platform
   http://www.ohse.de/uwe/articles/gcc-attributes.html#func-fastcall
   http://www.angelcode.com/dev/callconv/callconv.html#thiscall

[10] Borland register calling convention
    http://docwiki.embarcadero.com/RADStudio/en/Program_Control#Register_Convention

    http://www.openwatcom.org/index.php/Calling_Conventions

[12] Microsoft calling conventions / Calling conventions on the x86 platform
    http://www.angelcode.com/dev/callconv/callconv.html#thiscall

[13] Calling conventions on the x86 platform
    http://www.angelcode.com/dev/callconv/callconv.html#thiscall

[14] Pascal calling convention
    http://en.wikipedia.org/wiki/X86_calling_conventions#pascal
[15] Plan9 C compiler calling convention
http://www.mail-archive.com/9fans@9fans.net/msg16421.html

[16] ARM-THUMB Procedure Call Standard
http://tinyurl.com/2rxb3a

[17] Procedure Call Standard for the ARM Architecture

[18] Debian ARM EABI Port Wiki
http://wiki.debian.org/ArmEabiPort

[19] MSDN: x64 Software Conventions
http://tinyurl.com/2k3tfw

[20] System V Application Binary Interface - AMD64 Architecture Processor Supplement
http://tinyurl.com/2j5tex

[21] System V Application Binary Interface - SPARC Processor Supplement

[22] Introduction to Mac OS X ABI Function Call Guide
http://tinyurl.com/s7f85

[23] Linux Standard Base Core Specification for PPC32 3.2 - Chapter 8. Low Level System Information
http://tinyurl.com/9ttwcy


[25] Developing PowerPC Embedded Application Binary Interface (EABI)

[26] devkitPro - homebrew game development
http://www.devkitpro.org/

[27] psptoolchain - all the homebrew related material ps2dev.org
http://ps2dev.org/psp/

[28] a GEM Dynamical Library system for TOS computer
http://ldg.sourceforge.net/

[29] libffi - a portable foreign function interface library
http://sources.redhat.com/libffi/

[30] C/Invoke - library for connecting to C libraries at runtime
http://www.nongnu.org/cinvoke/

[31] libffcall - foreign function call libraries
http://www.haible.de/bruno/packages-ffcall.html
