



ClamAV Bytecode Compiler
User Manual

Contents

1	Installation	1
1.1	Requirements	1
1.2	Obtaining the ClamAV Bytecode Compiler	1
1.3	Building	2
1.3.1	Disk space	2
1.3.2	Create build directory	2
1.4	Testing	2
1.5	Installing	3
1.5.1	Structure of installed files	3
2	Tutorial	5
2.1	Short introduction to the bytecode language	5
2.1.1	Types, variables and constants	5
2.1.2	Arrays and pointers	5
2.1.3	Arithmetics	5
2.1.4	Functions	5
2.1.5	Control flow	5
2.1.6	Common functions	5
2.2	Writing logical signatures	5
2.2.1	Structure of a bytecode for algorithmic detection	5
2.2.2	Virusnames	6
2.2.3	Patterns	6
2.2.4	Single subsignature	8
2.2.5	Multiple subsignatures	9
2.2.6	W32.Polipos.A detector rewritten as bytecode	9
2.2.7	Virut detector in bytecode	9
2.3	Writing regular expressions in bytecode	9
2.3.1	A very simple regular expression	9
2.3.2	Named regular expressions	12
2.4	Writing unpackers	12
2.4.1	Structure of a bytecode for unpacking (and other hooks)	12
2.4.2	Detecting clam.exe via bytecode	13
2.4.3	Detecting clam.exe via bytecode (disasm)	13
2.4.4	A simple unpacker	13
2.4.5	Matching PDF javascript	13
2.4.6	YC unpacker rewritten as bytecode	13

3	Usage	15
3.1	Invoking the compiler	15
3.1.1	Compiling C++ files	15
3.2	Running compiled bytecode	16
3.2.1	ClamBC	16
3.2.2	clamscan, clamd	16
3.3	Debugging bytecode	16
3.3.1	“printf” style debugging	16
3.3.2	Single-stepping	17
4	ClamAV bytecode language	19
4.1	Differences from C99 and GNU C	19
4.2	Limitations	21
4.3	Logical signatures	21
4.4	Headers and runtime environment	23
5	Bytecode security & portability	25
6	Reporting bugs	27
7	Bytecode API	29
7.1	API groups	29
7.1.1	Bytecode configuration	29
7.1.2	Data structure handling functions	30
7.1.3	Disassemble APIs	32
7.1.4	Engine queries	32
7.1.5	Environment detection functions	32
7.1.6	File operations	33
7.1.7	Global variables	33
7.1.8	Icon matcher APIs	33
7.1.9	JS normalize API	34
7.1.10	Math functions	34
7.1.11	PDF handling functions	34
7.1.12	PE functions	35
7.1.13	Scan control functions	38
7.1.14	String operations	38
7.2	Structure types	39
7.2.1	cli_exe_info Struct Reference	39
7.2.1.1	Detailed Description	39
7.2.1.2	Field Documentation	39
7.2.2	cli_exe_section Struct Reference	39
7.2.2.1	Detailed Description	40
7.2.2.2	Field Documentation	40
7.2.3	cli_pe_hook_data Struct Reference	40
7.2.3.1	Detailed Description	41
7.2.3.2	Field Documentation	41
7.2.4	DIS_arg Struct Reference	41
7.2.4.1	Detailed Description	42
7.2.4.2	Field Documentation	42
7.2.5	DIS_fixed Struct Reference	42

7.2.5.1	Detailed Description	42
7.2.5.2	Field Documentation	42
7.2.6	DIS_mem_arg Struct Reference	43
7.2.6.1	Detailed Description	43
7.2.6.2	Field Documentation	43
7.2.7	DISASM_RESULT Struct Reference	43
7.2.7.1	Detailed Description	43
7.2.8	pe_image_data_dir Struct Reference	43
7.2.8.1	Detailed Description	43
7.2.9	pe_image_file_hdr Struct Reference	44
7.2.9.1	Detailed Description	44
7.2.9.2	Field Documentation	44
7.2.10	pe_image_optional_hdr32 Struct Reference	44
7.2.10.1	Detailed Description	45
7.2.10.2	Field Documentation	45
7.2.11	pe_image_optional_hdr64 Struct Reference	46
7.2.11.1	Detailed Description	46
7.2.11.2	Field Documentation	46
7.2.12	pe_image_section_hdr Struct Reference	47
7.2.12.1	Detailed Description	47
7.2.12.2	Field Documentation	47
7.3	Low level API	48
7.3.1	bytecode_api.h File Reference	48
7.3.1.1	Detailed Description	50
7.3.1.2	Enumeration Type Documentation	50
7.3.1.3	Function Documentation	51
7.3.1.4	Variable Documentation	72
7.3.2	bytecode_disasm.h File Reference	73
7.3.2.1	Detailed Description	75
7.3.2.2	Enumeration Type Documentation	75
7.3.3	bytecode_execs.h File Reference	83
7.3.3.1	Detailed Description	83
7.3.4	bytecode_pe.h File Reference	83
7.3.4.1	Detailed Description	84
7.4	High level API	84
7.4.1	bytecode_local.h File Reference	84
7.4.1.1	Detailed Description	86
7.4.1.2	Define Documentation	86
7.4.1.3	Function Documentation	89
7.4.1.4	Variable Documentation	105
8	Copyright and License	107
8.1	The ClamAV Bytecode Compiler	107
8.2	Bytecode	109

A Predefined macros**111**

ClamAV Bytecode Compiler - Internals Manual,

© 2009 Sourcefire, Inc.

Authors: Török Edvin

This document is distributed under the terms of the GNU General Public License v2.

Clam AntiVirus is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation; version 2 of the License.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301, USA.

ClamAV and Clam AntiVirus are trademarks of Sourcefire, Inc.

CHAPTER 1

Installation

1.1. Requirements

The ClamAV Bytecode Compiler uses the LLVM compiler framework, thus requires an Operating System where building LLVM is supported:

- FreeBSD/x86
- Linux/{x86,x86_64,ppc}
- Mac OS X/{x86,ppc}
- Solaris/sparcv9
- Windows/x86 using mingw32 or Visual Studio

The following packages are required to compile the ClamAV Bytecode Compiler:

- GCC C and C++ compilers (minimum 4.1.3, recommended: 4.3.4 or newer)¹.
- Perl (version 5.6.0+)
- GNU make (version 3.79+, recommended 3.81)

The following packages are optional, but highly recommended:

- Python (version 2.5.4+?) - for running the tests

1.2. Obtaining the ClamAV Bytecode Compiler

You can obtain the source code in one of the following ways ²

¹Note that several versions of GCC have bugs when compiling LLVM, see <http://llvm.org/docs/GettingStarted.html#brokengcc> for a full list. Also LLVM requires support for atomic builtins for multithreaded mode, which gcc 3.4.x doesn't have

²For now the use the internal clamtools repository:
git clone username@git.clam.sourcefire.com:/var/lib/git/clamtools.git

- Check out the source code using git native protocol:

```
git clone git://git.clamav.net/git/clamav-bytecode-compiler
```
 - Check out the source code using HTTP:

```
git clone http://git.clamav.net/git/clamav-bytecode-compiler.git
```
- You can keep the source code updated using:

```
git pull
```

1.3. Building

1.3.1. Disk space

A minimalistic release build requires 100M of disk space.

Testing the compiler requires a full build, 320M of disk space. A debug build requires significantly more disk space (1.4G for a minimalistic debug build).

Note that this is only needed during the build process, once installed only 12M is needed.

1.3.2. Create build directory

Building requires a separate object directory, building in the source directory is not supported. Create a build directory:

```
$ cd clamav-bytecode-compiler && mkdir obj
```

Run configure (you can use any prefix you want, this example uses /usr/local/clamav):

```
$ cd obj && ../llvm/configure --enable-optimized \
--enable-targets=host-only --disable-bindings \
--prefix=/usr/local/clamav
```

Run the build under ulimit ¹:

```
$ (ulimit -t 3600 -v 512000 && make clambc-only -j4)
```

1.4. Testing

```
$ (ulimit -t 3600 -v 512000 && make -j4)
$ make check-all
```

If make check reports errors, check that your compiler is NOT on this list: <http://llvm.org/docs/GettingStarted.html#brokengcc>.

If it is, then your compiler is buggy, and you need to do one of the following: upgrade your compiler to a non-buggy version, upgrade the OS to one that has a non-buggy compiler, compile with `export OPTIMIZE_OPTION=-O2`, or `export OPTIMIZE_OPTION=-O1`, or `export OPTIMIZE_OPTION=-O1`.

If not you probably found a bug, report it at <http://bugs.clamav.net>

¹compiling some files can be very memory intensive, especially with older compilers

1.5. Installing

Install it:

```
$ make install-clambc -j8
```

1.5.1. Structure of installed files

1. The ClamAV Bytecode compiler driver: `$PREFIX/bin/clambc-compiler`
2. ClamAV bytecode header files:

```
$PREFIX/lib/clang/1.1/include:  
bcfeatures.h  
bytecode_{api,decl,c,api,disasm,execs,features}.h  
bytecode.h  
bytecode_{local,pe,types}.h
```

3. clang compiler (with ClamAV bytecode backend) compiler include files:

```
$PREFIX/lib/clang/1.1/include:  
emmintrin.h  
float.h  
iso646.h  
limits.h  
{,p,t,x}mmmintrin.h  
mm_malloc.h  
std{arg,bool,def,int}.h  
tgmath.h
```

4. User manual

```
$PREFIX/docs/clamav/clambc-user.pdf
```


CHAPTER 2

Tutorial

2.1. Short introduction to the bytecode language

2.1.1. Types, variables and constants

2.1.2. Arrays and pointers

2.1.3. Arithmetics

2.1.4. Functions

2.1.5. Control flow

2.1.6. Common functions

2.2. Writing logical signature bytecodes

¹ Logical signatures can be used as triggers for executing bytecode. However, instead of describing a logical signature as a `.ldb` pattern, you use (simple) C code which is later translated to a `.ldb`-style logical signature by the ClamAV Bytecode Compiler.

A bytecode triggered by a logical signature is much more powerful than a logical signature itself: you can write complex algorithmic detections, and use the logical signature as a *filter* (to speed up matching). Thus another name for “logical signature bytecodes” is “algorithmic detection bytecodes”. The detection you write in bytecode has read-only access to the file being scanned and its metadata (PE sections, EP, etc.).

2.2.1. Structure of a bytecode for algorithmic detection

Algorithmic detection bytecodes are triggered when a logical signature matches. They can execute an algorithm that determines whether the file is infected and with which virus.

A bytecode can be either algorithmic or an unpacker (or other hook), but not both.

It consists of:

¹See Section 4.3 for more details about logical signatures in bytecode.

- Definition of virusnames used in the bytecode
- Pattern definitions (for logical subexpressions)
- The logical signature as C function: `bool logical_trigger(void)`
- The `int entrypoint(void)` function which gets executed when the logical signature matches
- (Optional) Other functions and global constants used in `entrypoint`

The syntax for defining logical signatures, and an example is described in Section 2.2.4.

The function `entrypoint` must report the detected virus by calling `foundVirus` and returning 0. It is recommended that you always return 0, otherwise a warning is shown and the file is considered clean. If `foundVirus` is not called, then ClamAV also assumes the file is clean.

2.2.2. Virusnames

Each logical signature bytecode must have a virusname prefix, and one or more virusnames. The virusname prefix is used by the SI to ensure unique virusnames (a unique number is appended for duplicate prefixes).

Program 1 Declaring virusnames

```

1 /* Prefix, used for duplicate detection and fixing */
  VIRUSNAME_PREFIX("Trojan.Foo")
3 /* You are only allowed to set these virusnames as found */
  VIRUSNAMES("A", "B")
5 /* File type */
  TARGET(2)
```

In Program 1 3 predefined macros are used:

- `VIRUSNAME_PREFIX` which must have exactly one string argument
- `VIRUSNAMES` which must have one or more string arguments
- `TARGET` which must have exactly one integer argument

In this example, the bytecode could generate one of these virusnames: `Trojan.Foo.A`, or `Trojan.Foo.B`, by calling `foundVirus("A")` or `foundVirus("B")` respectively (notice that the prefix is not part of these calls).

2.2.3. Patterns

Logical signatures use `.ndb` style patterns, an example on how to define these is shown in Program 2.

Each pattern has a name (like a variable), and a string that is the hex pattern itself. The declarations are delimited by the macros `SIGNATURES_DECL_BEGIN`, and `SIGNATURES_DECL_END`. The definitions are delimited by the macros `SIGNATURES_DEF_BEGIN`, and `SIGNATURES_END`. Declarations must always come

Program 2 Declaring patterns

```

SIGNATURES_DECL_BEGIN
2 DECLARE_SIGNATURE(magic)
  DECLARE_SIGNATURE(check)
4 DECLARE_SIGNATURE(zero)
  SIGNATURES_DECL_END
6
SIGNATURES_DEF_BEGIN
8 DEFINE_SIGNATURE(magic, "EP+0:aabb")
  DEFINE_SIGNATURE(check, "f00d")
10 DEFINE_SIGNATURE(zero, "fff")
  SIGNATURES_END

```

before definitions, and you can have only one declaration and declaration section! (think of declaration like variable declarations, and definitions as variable assignments, since that's what they are under the hood). The order in which you declare the signatures is the order in which they appear in the generated logical signature.

You can use any name for the patterns that is a valid record field name in C, and doesn't conflict with anything else declared.

After using the above macros, the global variable `Signatures` will have two new fields: `magic`, and `zero`. These can be used as arguments to the functions `count_match()`, and `matches()` anywhere in the program as shown in Program 3:

- `matches(Signatures.match)` will return true when the `match` signature matches (at least once)
- `count_match(Signatures.zero)` will return the number of times the `zero` signature matched
- `count_match(Signatures.check)` will return the number of times the `check` signature matched

The condition in the `if` can be interpreted as: if the `match` signature has matched at least once, and the number of times the `zero` signature matched is higher than the number of times the `check` signature matched, then we have found a virus A, otherwise the file is clean.

Program 3 Using patterns

```

1 int entrypoint(void)
  {
3   if (matches(Signatures.match) && count_match(Signatures.zero) >
      count_match(Signatures.check))
      foundVirus("A");
5   return 0;
  }

```

2.2.4. Single subsignature

The simplest logical signature is like a `.ndb` signature: a virus name, signature target, 0 as logical expression ¹, and a `ndb`-style pattern.

The code for this is shown in Program 4

Program 4 Single subsignature example

```

/* Declare the prefix of the virusname */
2 VIRUSNAME_PREFIX("Trojan.Foo")
/* Declare the suffix of the virusname */
4 VIRUSNAMES("A")
/* Declare the signature target type (1 = PE) */
6 TARGET(1)

8 /* Declare the name of all subsignatures used */
SIGNATURES_DECL_BEGIN
10 DECLARE_SIGNATURE(magic)
SIGNATURES_DECL_END
12
/* Define the pattern for each subsignature */
14 SIGNATURES_DEF_BEGIN
DEFINE_SIGNATURE(magic, "aabb")
16 SIGNATURES_END

18 /* All bytecode triggered by logical signatures must have this
   function */
20 bool logical_trigger(void)
{
22     /* return true if the magic subsignature matched,
       * its pattern is defined above to "aabb" */
24     return count_match(Signatures.magic) != 2;
}

26
/* This is the bytecode function that is actually executed when
   the logical
28 * signature matched */
int entrypoint(void)
30 {
    /* call this function to set the suffix of the virus found */
32     foundVirus("A");
    /* success, return 0 */
34     return 0;
}

```

The logical signature (created by the compiler) looks like this:
`Trojan.Foo.{A};Target:2;0;aabb`

Of course you should use a `.ldb` signature in this case when all the processing in `entrypoint` is only setting a virusname and returning. However, you can do more complex checks in `entrypoint`, once the bytecode was triggered by the `logical_trigger`

In the example in Program 4 the pattern was used without an anchor; such a pattern matches at any offset. You can use offsets though, the same way as in `.ndb` signatures, see Program 5 for an example.

¹meaning that subexpression 0 must match

2.2.5. Multiple subsignatures

An example for this is shown in Program 5. Here you see the following new features used: ¹

- Multiple virusnames returned from a single bytecode (with common prefix)
- Multiple subsignatures, each with a name of your choice
- A pattern with an anchor (EP+0:aabb)
- More subsignatures defined than used in the logical expression

The logical signature looks like this:

```
Trojan.Foo.{A,B};Target:2;(((0|1|2)=42,2)|(3=10));EP+0:aabb;ffff;aaccee;f00d;dead
```

Notice how the subsignature that is not used in the logical expression (number 4, `dead`) is used in `entrypoint` to decide the virus name. This works because ClamAV does collect the match counts for all subsignatures (regardless if they are used or not in a signature). The `count_match(Signatures.check2)` call is thus a simple memory read of the count already determined by ClamAV.

Also notice that comments can be used freely: they are ignored by the compiler. You can use either C-style multiline comments (start comment with `/*`, end with `*/`), or C++-style single-line comments (start comment with `//`, automatically ended by newline).

2.2.6. W32.Polipos.A detector rewritten as bytecode

2.2.7. Virut detector in bytecode

2.3. Writing regular expressions in bytecode

ClamAV only supports a limited set of regular expressions in `.ndb format`: wildcards. The bytecode compiler allows you to compile fully generic regular expressions to bytecode directly. When `libclamav` loads the bytecode, it will compile to native code (if using the JIT), so it should offer quite good performance.

The compiler currently uses `re2c` to compile regular expressions to C code, and then compile that to bytecode. The internal workings are all transparent to the user: the compiler automatically uses `re2c` when needed, and `re2c` is embedded in the compiler, so you don't need to install it.

The syntax of regular expressions are similar to the one used by POSIX regular expressions, except you have to quote literals, since unquoted they are interpreted as regular expression names.

2.3.1. A very simple regular expression

Lets start with a simple example, to match this POSIX regular expression: `eval([a-zA-Z_][a-zA-Z0-9_]*\.`unescape.

See Program 6 ².

There are several new features introduced here, here is a step by step breakdown:

¹In case of a duplicate virusname the prefix is appended a unique number by the SI

²This omits the virusname, and logical signature declarations

Program 5 Multiple subsignatures

```

1  /* You are only allowed to set these virusnames as found */
   VIRUSNAME_PREFIX("Test")
3  VIRUSNAMES("A", "B")
   TARGET(1)
5
   SIGNATURES_DECL_BEGIN
7  DECLARE_SIGNATURE(magic)
   DECLARE_SIGNATURE(zero)
9  DECLARE_SIGNATURE(check)
   DECLARE_SIGNATURE(fivetoten)
11 DECLARE_SIGNATURE(check2)
   SIGNATURES_DECL_END
13
   SIGNATURES_DEF_BEGIN
15 DEFINE_SIGNATURE(magic, "EP+0:aabb")
   DEFINE_SIGNATURE(zero, "ffff")
17 DEFINE_SIGNATURE(fivetoten, "aaccee")
   DEFINE_SIGNATURE(check, "f00d")
19 DEFINE_SIGNATURE(check2, "dead")
   SIGNATURES_END
21
   bool logical_trigger(void)
23 {
       unsigned sum_matches = count_match(Signatures.magic)+
25         count_match(Signatures.zero) +
           count_match(Signatures.fivetoten);
       unsigned unique_matches = matches(Signatures.magic)+
27         matches(Signatures.zero)+
           matches(Signatures.fivetoten);
       if (sum_matches == 42 && unique_matches == 2) {
29         // The above 3 signatures have matched a total of 42
           times, and at least
           // 2 of them have matched
31         return true;
       }
33     // If the check signature matches 10 times we still have a
       match
       if (count_match(Signatures.check) == 10)
35         return true;
       // No match
37     return false;
   }
39
   int entrypoint(void)
41 {
       unsigned count = count_match(Signatures.check2);
43       if (count >= 2)
           foundVirus(count == 2 ? "A" : "B");
       //
45       if (count == 2)
           foundVirus("A");
47       else
           foundVirus("B");
49       return 0;
   }

```

Program 6 Simple regular expression example

```

1 int entrypoint(void)
2 {
3     REGEX_SCANNER;
4     seek(0, SEEK_SET);
5     for (;;) {
6         REGEX_LOOP_BEGIN
7
8         /*!re2c
9          ANY = [^];
10
11         "eval("[a-zA-Z_][a-zA-Z_0-9]*".unescape" {
12             long pos = REGEX_POS;
13             if (pos < 0)
14                 continue;
15             debug("unescape found at");
16             debug(pos);
17         }
18         ANY { continue; }
19     */
20 }
21 return 0;
22 }

```

`REGEX_SCANNER` this declares the data structures needed by the regular expression matcher

`seek(0, SEEK_SET)` this sets the current file offset to position 0, matching will start at this position. For offset 0 it is not strictly necessary to do this, but it serves as a reminder that you might want to start matching somewhere, that is not necessarily 0.

`for(;;) { REGEX_LOOP_BEGIN` this creates the regular expression matcher main loop. It takes the current file byte-by-byte ¹ and tries to match one of the regular expressions.

`/*!re2c` This marks the beginning of the regular expression description. The entire regular expression block is a C comment, starting with `!re2c`

`ANY = [^];` This declares a regular expression named `ANY` that matches any byte.

`"eval("[a-zA-Z_][a-zA-Z_0-9]*".unescape" {` This is the actual regular expression.

`"eval("` This matches the literal string `eval(`. Literals have to be placed in double quotes `"` here, unlike in POSIX regular expressions or PCRE. If you want case-insensitive matching, you can use `'`.

`[a-zA-Z_]` This is a character class, it matches any lowercase, uppercase or `_` characters.

`[a-zA-Z_0-9]*` Same as before, but with repetition. `*` means match zero or more times, `+` means match one or more times, just like in POSIX regular expressions.

¹it is not really reading byte-by-byte, it is using a buffer to speed things up


```

    ".unescape" A literal string again
    { start of the action block for this regular expression. Whenever the
      regular expression matches, the attached C code is executed.

long pos = REGEX_POS; this determines the absolute file offset where the
regular expression has matched. Note that because the regular expression
matcher uses a buffer, using just seek(0, SEEK_CUR) would give the
current position of the end of that buffer, and not the current position
during regular expression matching. You have to use the REGEX_POS macro
to get the correct position.

debug(...) Shows a debug message about what was found and where. This is
extremely helpful when you start writing regular expressions, and nothing
works: you can determine whether your regular expression matched
at all, and if it matched where you thought it would. There is also a
DEBUG_PRINT_MATCH that prints the entire matched string to the debug
output. Of course before publishing the bytecode you might want to turn
off these debug messages.

} closes the action block for this regular expression

ANY { continue; } If none of the regular expressions matched so far, just
keep running the matcher, at the next byte

*/ closes the regular expression description block

} closes the for() loop

```

You may have multiple regular expressions, or declare multiple regular expressions with a name, and use those names to build more complex regular expressions.

2.3.2. Named regular expressions

2.4. Writing unpackers

2.4.1. Structure of a bytecode for unpacking (and other hooks)

When writing an unpacker, the bytecode should consist of:

- Define which hook you use (for example `PE_UNPACKER_DECLARE` for a PE hook)
- An `int entrypoint(void)` function that reads the current file and unpacks it to a new file
- Return 0 from `entrypoint` if you want the unpacked file to be scanned
- (Optional) Other functions and global constants used by `entrypoint`

2.4.2. Detecting clam.exe via bytecode

Example provided by aCaB:

2.4.3. Detecting clam.exe via bytecode (disasm)

Example provided by aCaB:

2.4.4. A simple unpacker

2.4.5. Matching PDF javascript

2.4.6. YC unpacker rewritten as bytecode

CHAPTER 3

Usage

3.1. Invoking the compiler

Compiling is similar to gcc ¹:

```
$ /usr/local/clamav/bin/clambc-compiler foo.c -o foo.cbc -O2
```

This will compile the file `foo.c` into a file called `foo.cbc`, that can be loaded by ClamAV, and packed inside a `.cvd` file.

The compiler by default has all warnings turned on.

Supported optimization levels: `-O0`, `-O1`, `-O2`, `-O3`. ² It is recommended that you always compile with at least `-O1`.

Warning options: `-Werror` (transforms all warnings into errors).

Preprocessor flags:

-I <directory> Searches in the given directory when it encounters a `#include "headerfile"` directive in the source code, in addition to the system defined header search directories.

-D <MACRONAME>=<VALUE> Predefine given `<MACRONAME>` to be equal to `<VALUE>`.

-U <MACRONAME> Undefine a predefined macro

The compiler also supports some other commandline options (see `clambc-compiler --help` for a full list), however some of them have no effect when using the ClamAV bytecode backend (such as the X86 backend options). You shouldn't need to use any flags not documented above.

3.1.1. Compiling C++ files

Filenames with a `.cpp` extension are compiled as C++ files, however `clang++` is not yet ready for production use, so this is EXPERIMENTAL currently. For now write bytecodes in C.

¹Note that the ClamAV bytecode compiler will refuse to compile code it considers insecure

²Currently `-O0` doesn't work

3.2. Running compiled bytecode

After compiling a C source file to bytecode, you can load it in ClamAV:

3.2.1. ClamBC

ClamBC is a tool you can use to test whether the bytecode loads, compiles, and can execute its entrypoint successfully. Usage:

```
clambc <file> [function] [param1 ...]
```

For example loading a simple bytecode with 2 functions is done like this:

```
$ clambc foo.cbc
LibClamAV debug: searching for unrar, user-searchpath: /usr/local/lib
LibClamAV debug: unrar support loaded from libclamunrar_iface.so.6.0.4 libclamunrar_i
LibClamAV debug: bytecode: Parsed 0 APICalls, maxapi 0
LibClamAV debug: Parsed 1 BBs, 2 instructions
LibClamAV debug: Parsed 1 BBs, 2 instructions
LibClamAV debug: Parsed 2 functions
Bytecode loaded
Running bytecode function :0
Bytecode run finished
Bytecode returned: 0x8
Exiting
```

3.2.2. clamscan, clamd

You can tell clamscan to load the bytecode as a database directly:

```
$ clamscan -dfoo.cbc
```

Or you can instruct it to load all databases from a directory, then clamscan will load all supported formats, including files with bytecode, which have the .cbc extension.

```
$ clamscan -ddirectory
```

You can also put the bytecode files into the default database directory of ClamAV (usually /usr/local/share/clamav) to have it loaded automatically from there. Of course, the bytecode can be stored inside CVD files, too.

3.3. Debugging bytecode

3.3.1. “printf” style debugging

Printf, and printf-like format specifiers are not supported in the bytecode. You can use these functions instead of printf to print strings and integer to clamscan’s -debug output:

`debug_print_str`, `debug_print_uint`, `debug_print_str_start`, `debug_print_str_nonl`.

You can also use the `debug` convenience wrapper that automatically prints as string or integer depending on parameter type: `debug`, `debug`, `debug`.

See Program 7 for an example.

Program 7 Example of using debug APIs

```

/* test debug APIs */
2 int entrypoint(void)
{
4     /* print a debug message, followed by newline */
    debug_print_str("bytecode started", 16);

6
    /* start a new debug message, don't end with newline yet */
8     debug_print_str_start("Engine functionality level: ", 28);
    /* print an integer, no newline */
10    debug_print_uint(engine_functionality_level());
    /* print a string without starting a new debug message, and
        without
12    * terminating with newline */
    debug_print_str_nonl(" dconf functionality level: ", 28);
14    debug_print_uint(engine_dconf_level());
    debug_print_str_nonl("\n", 1);
16    debug_print_str_start("Engine scan options: ", 21);
    debug_print_uint(engine_scan_options());
18    debug_print_str_nonl(" db options: ", 13);
    debug_print_uint(engine_db_options());
20    debug_print_str_nonl("\n", 1);

22    /* convenience wrapper to just print a string */
    debug("just print a string");
24    /* convenience wrapper to just print an integer */
    debug(4);
26    return 0xf00d;
}

```

3.3.2. Single-stepping

If you have GDB 7.0 (or newer) you can single-step ¹ ² during the execution of the bytecode.

- Run clambc or clamscan under gdb:

```

$ ./libtool --mode=execute gdb clamscan/clamscan
...
(gdb) b cli_vm_execute_jit
Are you sure ....? y
(gdb) run -dfoo.cbc
...
Breakpoint ....

(gdb) step
(gdb) next

```

You can single-step through the execution of the bytecode, however you can't (yet) print values of individual variables, you'll need to add debug statements in the bytecode to print interesting values.

¹not yet implemented in libclamav

²assuming you have JIT support

CHAPTER 4

ClamAV bytecode language

The bytecode that ClamAV loads is a simplified form of the LLVM Intermediate Representation, and as such it is language-independent.

However currently the only supported language from which such bytecode can be generated is a simplified form of C ¹

The language supported by the ClamAV bytecode compiler is a restricted set of C99 with some GNU extensions.

4.1. Differences from C99 and GNU C

These restrictions are enforced at compile time:

- No standard include files. ²
- The ClamAV API header files are preincluded.
- No external function calls, except to the ClamAV API ³
- No inline assembly ⁴
- Globals can only be readonly constants ⁵
- `inline` is C99 inline (equivalent to GNU C89 extern inline), thus it cannot be used outside of the definition of the ClamAV API, you should use `static inline`
- `sizeof(int) == 4` always
- `sizeof(long) == sizeof(long long) == 8` always
- `ptrdiff_t = int,` `intptr_t = int,` `intmax_t = long,`
 `uintmax_t = unsigned long` ⁶

¹In the future more languages could be supported, see the Internals Manual on language frontends

²For portability reasons: preprocessed C code is not portable

³For safety reasons we can't allow the bytecode to call arbitrary system functions

⁴This is both for safety and portability reasons

⁵For thread safety reasons

⁶Note that a pointer's `sizeof` is runtime-platform dependent, although at compile time `sizeof(void*) == 4`, at runtime it can be something else. Thus you should avoid using `sizeof(pointer)`

- No pointer to integer casts and integer to pointer casts (pointer arithmetic is allowed though)
- No `__thread` support
- Size of memory region associated with each pointer must be known in each function, thus if you pass a pointer to a function, you must also pass its allocated size as a parameter.
- Endianness must be handled via the `__is_bigendian()` API function call, or via the `cli_{read,write}int{16,32}` wrappers, and not by casting pointers
- Predefines `__CLAMBC__`
- All integer types have fixed width
- `main` or `entrypoint` must have the following prototype: `int main(void)`, the prototype `int main(int argc, char *argv[])` is not accepted

They are meant to ensure the following:

- Thread safe execution of multiple different bytecodes, and multiple instances of the same bytecode
- Portability to multiple CPU architectures and OSes: the bytecode must execute on both the libclamav/LLVM JIT where that is supported (x86, x86_64, ppc, arm?), and on the libclamav interpreter where that is not supported.
- No external runtime dependency: libclamav should have everything needed to run the bytecode, thus no external calls are allowed, not even to libc!
- Same behaviour on all platforms: fixed size integers.

These restrictions are checked at runtime (checks are inserted at compile time):

- Accessing an out-of-bounds pointer will result in a call to `abort()`
- Calling `abort()` interrupts the execution of the bytecode in a thread safe manner, and doesn't halt ClamAV ¹.

The ClamAV API header has further restriction, see the Internals manual.

Although the bytecode undergoes a series of automated tests (see Publishing chapter in Internals manual), the above restrictions don't guarantee that the resulting bytecode will execute correctly! You must still test the code yourself, these restrictions only avoid the most common errors. Although the compiler and verifier aims to accept only code that won't crash ClamAV, no code is 100% perfect, and a bug in the verifier could allow unsafe code be executed by ClamAV.

¹in fact it calls a ClamAV API function, and not the libc abort function.

4.2. Limitations

The bytecode format has the following limitations:

- At most 64k bytecode kinds (hooks)
- At most 64k types (including pointers, and all nested types)
- At most 16 parameters to functions, no vararg functions
- At most 64-bit integers
- No vector types or vector operations
- No opaque types
- No floating point
- Global variable initializer must be compile-time computable
- At most 32k global variables (and at most 32k API globals)
- Pointer indexing at most 15 levels deep (can be worked around if needed by using temporaries)
- No struct return or byval parameters
- At most 32k instructions in a single function
- No Variable Length Arrays

4.3. Logical signatures

Logical signatures can be used as triggers for executing a bytecode. Instead of describing a logical signatures as a `.ldb` pattern, you use C code which is then translated to a `.ldb`-style logical signature.

Logical signatures in ClamAV support the following operations:

- Sum the count of logical subsignatures that matched inside a subexpression
- Sum the number of different subsignatures that matched inside a subexpression
- Compare the above counts using the `>`, `=`, `<` relation operators
- Perform logical `&&`, `||` operations on above boolean values
- Nest subexpressions
- Maximum 64 subexpressions

Out of the above operations the ClamAV Bytecode Compiler doesn't support computing sums of nested subexpressions, (it does support nesting though).

The C code that can be converted into a logical signature must obey these restrictions:

- a function named `logical_trigger` with the following prototype:
`bool logical_trigger(void)`
- no function calls, except for `count_match` and `matches`
- no global variable access (except as done by the above 2 functions internally)
- return true when signature should trigger, false otherwise
- use only integer compare instructions, branches, integer *add*, logical *and*, logical *or*, logical *xor*, zero extension, store/load from local variables
- the final boolean expression must be convertible to disjunctive normal form without negation
- the final logical expression must not have more than 64 subexpressions
- it can have early returns (all true returns are unified using `||`)
- you can freely use comments, they are ignored
- the final boolean expression cannot be a `true` or `false` constant

The compiler does the following transformations (not necessarily in this order):

- convert shortcircuit boolean operations into non-shortcircuit ones (since all operands are boolean expressions or local variables, it is safe to execute these unconditionally)
- propagate constants
- simplify control flow graph
- (sparse) conditional constant propagation
- dead store elimination
- dead code elimination
- instruction combining (arithmetic simplifications)
- jump threading

If after this transformation the program meets the requirements outlined above, then it is converted to a logical signature. The resulting logical signature is simplified using basic properties of boolean operations, such as associativity, distributivity, De Morgan's law.

The final logical signature is not unique (there might be another logical signature with identical behavior), however the boolean part is in a canonical form: it is in disjunctive normal form, with operands sorted in ascending order.

For best results the C code should consist of:

- local variables declaring the sums you want to use
- a series of `if` branches that `return true`, where the `if`'s condition is a single comparison or a logical *and* of comparisons

- a final `return false`

You can use `||` in the `if` condition too, but be careful that after expanding to disjunctive normal form, the number of subexpressions doesn't exceed 64.

Note that you do not have to use all the subsignatures you declared in `logical_trigger`, you can do more complicated checks (that wouldn't obey the above restrictions) in the bytecode itself at runtime. The `logical_trigger` function is fully compiled into a logical signature, it won't be a runtime executed function (hence the restrictions).

4.4. Headers and runtime environment

When compiling a bytecode program, `bytecode.h` is automatically included, so you don't need to explicitly include it. These headers (and the compiler itself) predefine certain macros, see Appendix A for a full list. In addition the following types are defined:

```
typedef unsigned char uint8_t;
2 typedef char int8_t;
typedef unsigned short uint16_t;
4 typedef short int16_t;
typedef unsigned int uint32_t;
6 typedef int int32_t;
typedef unsigned long uint64_t;
8 typedef long int64_t;
typedef unsigned int size_t;
10 typedef int off_t;
typedef struct signature { unsigned id } __Signature;
```

As described in Section 4.1 the width of integer types are fixed, the above typedefs show that.

A bytecode's entrypoint is the function `entrypoint` and it's required by ClamAV to load the bytecode.

Bytecode that is triggered by a logical signature must have a list of virusnames and patterns defined. Bytecodes triggered via hooks can optionally have them, but for example a PE unpacker doesn't need virus names as it only processes the data.

CHAPTER 5

Bytecode security & portability

CHAPTER 6

Reporting bugs

CHAPTER 7

Bytecode API

7.1. API groups

7.1.1. Bytecode configuration

Global **COPYRIGHT**(c) This will also prevent the sourcecode from being embedded into the bytecode

Global **DECLARE_SIGNATURE**(name)

Global **DEFINE_SIGNATURE**(name, hex)

Global **FUNCTIONALITY_LEVEL_MAX**(m)

Global **FUNCTIONALITY_LEVEL_MIN**(m)

Global **ICONGROUP1**(group)

Global **ICONGROUP2**(group)

Global **PDF_HOOK_DECLARE** This hook is called several times, use `pdf_get_phase()` to find out in which phase you got called.

Global **PE_UNPACKER_DECLARE**

Global **SIGNATURES_DECL_BEGIN**

Global **SIGNATURES_DECL_END**

Global **SIGNATURES_DEF_BEGIN**

Global **SIGNATURES_END**

Global **TARGET**(tgt)

Global **VIRUSNAME_PREFIX**(name)

Global **VIRUSNAMES**(...)

7.1.2. Data structure handling functions

Global `buffer_pipe_done(int32_t id)` After this all attempts to use this buffer will result in error. All `buffer_pipes` are automatically deallocated when bytecode finishes execution.

Global `buffer_pipe_new(uint32_t size)`

Global `buffer_pipe_new_fromfile(uint32_t pos)`

Global `buffer_pipe_read_avail(int32_t id)`

Global `buffer_pipe_read_get(int32_t id, uint32_t amount)` The 'amount' parameter should be obtained by a call to `buffer_pipe_read_avail()`.

Global `buffer_pipe_read_stopped(int32_t id, uint32_t amount)`
Updates read cursor in `buffer_pipe`.

Global `buffer_pipe_write_avail(int32_t id)`

Global `buffer_pipe_write_get(int32_t id, uint32_t size)` Returns pointer to writable buffer. The 'amount' parameter should be obtained by a call to `buffer_pipe_write_avail()`.

Global `buffer_pipe_write_stopped(int32_t id, uint32_t amount)`

Global `cli_readint16(const void *buff)`

Global `cli_readint32(const void *buff)`

Global `cli_writeint32(void *offset, uint32_t v)`

Global `hashset_add(int32_t hs, uint32_t key)`

Global `hashset_contains(int32_t hs, uint32_t key)`

Global `hashset_done(int32_t id)` Trying to use the hashset after this will result in an error. The hashset may not be used after this. All hashsets are automatically deallocated when bytecode finishes execution.

Global `hashset_empty(int32_t id)`

Global `hashset_new`(void)

Global `hashset_remove`(int32_t hs, uint32_t key)

Global `inflate_done`(int32_t id)

Global `inflate_init`(int32_t from_buffer, int32_t to_buffer, int32_t windowBits)
'from_buffer' and writing uncompressed data 'to_buffer'.

Global `inflate_process`(int32_t id)

Global `le16_to_host`(uint16_t v)

Global `le32_to_host`(uint32_t v)

Global `le64_to_host`(uint64_t v)

Global `malloc`(uint32_t size)

Global `map_addkey`(const uint8_t *key, int32_t ksize, int32_t id)

Global `map_done`(int32_t id)

Global `map_find`(const uint8_t *key, int32_t ksize, int32_t id)

Global `map_getvalue`(int32_t id, int32_t size)

Global `map_getvaluesize`(int32_t id)

Global `map_new`(int32_t keysize, int32_t valuesize)

Global `map_remove`(const uint8_t *key, int32_t ksize, int32_t id)

Global `map_setvalue`(const uint8_t *value, int32_t vsize, int32_t id)

7.1.3. Disassemble APIs

Global `__attribute__`

Class `DIS_arg`

Class `DIS_fixed`

Global `disasm_x86`(struct `DISASM_RESULT` *result, uint32_t len)

Global `DisassembleAt`(struct `DIS_fixed` *result, uint32_t offset, uint32_t len)

7.1.4. Engine queries

Global `count_match`(__Signature sig)

Global `engine_db_options`(void)

Global `engine_dconf_level`(void)

Global `engine_functionality_level`(void)

Global `engine_scan_options`(void)

Global `match_location`(__Signature sig, uint32_t goback)

Global `match_location_check`(__Signature sig, uint32_t goback, const char *static_sta
It is recommended to use this for safety and compatibility with 0.96.1

Global `matches`(__Signature sig)

7.1.5. Environment detection functions

Global `__is_bigendian`(void) `__attribute__((const))` `__attribute__((nothrow))`

Global `check_platform`(uint32_t a, uint32_t b, uint32_t c)

Global `disable_bytecode_if`(const int8_t *reason, uint32_t len, uint32_t cond)

Global `disable_jit_if`(const int8_t *reason, uint32_t len, uint32_t cond)

Global `get_environment`(struct cli_environment *env, uint32_t len)

Global `version_compare`(const uint8_t *lhs, uint32_t lhs_len, const uint8_t *rhs, uint32_t rhs_len)

7.1.6. File operations

Global `buffer_pipe_new_fromfile(uint32_t pos)` to the current file, at the specified position.

Global `file_byteat(uint32_t offset)`

Global `file_find(const uint8_t *data, uint32_t len)`

Global `file_find_limit(const uint8_t *data, uint32_t len, int32_t maxpos)`

Global `fill_buffer(uint8_t *buffer, uint32_t len, uint32_t filled, uint32_t cursor, uint32_t fill)`

Global `getFileSize(void)`

Global `read(uint8_t *data, int32_t size)`

Global `read_number(uint32_t radix)` Non-numeric characters are ignored.

Global `seek(int32_t pos, uint32_t whence)`

Global `write(uint8_t *data, int32_t size)`

7.1.7. Global variables

Global `__clambc_filesize[1]`

Global `__clambc_kind`

Global `__clambc_match_counts[64]`

Global `__clambc_match_offsets[64]`

Global `__clambc_pedata`

7.1.8. Icon matcher APIs

Global `matchicon(const uint8_t *group1, int32_t group1_len, const uint8_t *group2, int32_t group2_len)`

7.1.9. JS normalize API

Global `jsnorm_done`(int32_t id)

Global `jsnorm_init`(int32_t from_buffer)

Global `jsnorm_process`(int32_t id)

7.1.10. Math functions

Global `icos`(int32_t a, int32_t b, int32_t c)

Global `iexp`(int32_t a, int32_t b, int32_t c)

Global `ilog2`(uint32_t a, uint32_t b)

Global `ipow`(int32_t a, int32_t b, int32_t c)

Global `isin`(int32_t a, int32_t b, int32_t c)

7.1.11. PDF handling functions

Global `pdf_get_dumpedobjid`(void) Valid only in PDF_PHASE_POSTDUMP.

Global `pdf_get_flags`(void)

Global `pdf_get_obj_num`(void)

Global `pdf_get_phase`(void) Identifies at which phase this bytecode was called.

Global `pdf_getobj`(int32_t objidx, uint32_t amount) Meant only for reading, write modifies the fmap buffer, so avoid!

Global `pdf_getobjsize`(int32_t objidx)

Global `pdf_lookupobj`(uint32_t id)

Global `pdf_set_flags`(int32_t flags)

7.1.12. PE functions

Class `cli_exe_info`

Class `cli_exe_section`

Class `cli_pe_hook_data`

Global `get_pe_section`(struct `cli_exe_section` *section, uint32_t num)

Global `getEntryPoint`(void)

Global `getExeOffset`(void)

Global `getImageBase`(void)

Global `getNumberOfSections`(void)

Global `getPEBaseOfCode`(void)

Global `getPEBaseOfData`(void)

Global `getPECharacteristics`()

Global `getPEChecksum`(void)

Global `getPEDataDirRVA`(unsigned n)

Global `getPEDataDirSize`(unsigned n)

Global `getPEDllCharacteristics`(void)

Global `getPEFileAlignment`(void)

Global `getPEImageBase`(void)

Global `getPEisDLL`()

Global `getPELFANew`(void)

Global [getPELoaderFlags\(void\)](#)

Global [getPEMachine\(\)](#)

Global [getPEMajorImageVersion\(void\)](#)

Global [getPEMajorLinkerVersion\(void\)](#)

Global [getPEMajorOperatingSystemVersion\(void\)](#)

Global [getPEMajorSubsystemVersion\(void\)](#)

Global [getPEMinorImageVersion\(void\)](#)

Global [getPEMinorLinkerVersion\(void\)](#)

Global [getPEMinorOperatingSystemVersion\(void\)](#)

Global [getPEMinorSubsystemVersion\(void\)](#)

Global [getPENumberOfSymbols\(\)](#)

Global [getPEPointerToSymbolTable\(\)](#)

Global [getPESectionAlignment\(void\)](#)

Global [getPESizeOfCode\(void\)](#)

Global [getPESizeOfHeaders\(void\)](#)

Global [getPESizeOfHeapCommit\(void\)](#)

Global [getPESizeOfHeapReserve\(void\)](#)

Global [getPESizeOfImage\(void\)](#)

Global [getPESizeOfInitializedData\(void\)](#)

Global [getPESizeOfOptionalHeader\(\)](#)

Global `getPESizeOfStackCommit`(void)

Global `getPESizeOfStackReserve`(void)

Global `getPESizeOfUninitializedData`(void)

Global `getPESubsystem`(void)

Global `getPETimeDateStamp`()

Global `getPEWin32VersionValue`(void)

Global `getSectionRVA`(unsigned i) .

Global `getSectionVirtualSize`(unsigned i) .

Global `getVirtualEntryPoint`(void)

Global `hasExeInfo`(void)

Global `isPE64`(void)

Class `pe_image_data_dir`

Class `pe_image_file_hdr`

Class `pe_image_optional_hdr32`

Class `pe_image_optional_hdr64`

Class `pe_image_section_hdr`

Global `pe_rawaddr`(uint32_t rva)

Global `readPESectionName`(unsigned char name[8], unsigned n)

Global `readRVA`(uint32_t rva, void *buf, size_t bufsize)

7.1.13. Scan control functions

Global `bytecode_rt_error`(int32_t locationid)

Global `extract_new`(int32_t id)

Global `extract_set_container`(uint32_t container)

Global `foundVirus`(const char *virusname)

Global `input_switch`(int32_t extracted_file)

Global `setvirusname`(const uint8_t *name, uint32_t len)

7.1.14. String operations

Global `atoi`(const uint8_t *str, int32_t size)

Global `debug_print_str`(const uint8_t *str, uint32_t len)

Global `debug_print_str_nonl`(const uint8_t *str, uint32_t len)

Global `debug_print_str_start`(const uint8_t *str, uint32_t len)

Global `debug_print_uint`(uint32_t a)

Global `entropy_buffer`(uint8_t *buffer, int32_t size)

Global `hex2ui`(uint32_t hex1, uint32_t hex2)

Global `memchr`(const void *s, int c, size_t n)

Global `memcmp`(const void *s1, const void *s2, uint32_t n) __attribute__((__nothrow__))

Global `memcpy`(void *restrict dst, const void *restrict src, uintptr_t n) __attribute__((__nothrow__))

Global `memmove`(void *dst, const void *src, uintptr_t n) __attribute__((__nothrow__))

Global `memset`(void *src, int c, uintptr_t n) __attribute__((__nothrow__)) __attribute__((__nonnull__))

Global `memstr`(const uint8_t *haystack, int32_t haysize, const uint8_t *needle, int32_t n)

7.2. Structure types

7.2.1. cli_exe_info Struct Reference

Data Fields

- struct [cli_exe_section](#) * [section](#)
- uint32_t [offset](#)
- uint32_t [ep](#)
- uint16_t [nsections](#)
- struct [cli_hashset](#) * [vinfo](#)
- uint32_t [res_addr](#)
- uint32_t [hdr_size](#)

7.2.1.1. Detailed Description

Executable file information

[PE](#)

7.2.1.2. Field Documentation

- | | |
|--|---|
| 7.2.1.2.1. uint32_t ep | Entrypoint of executable |
| 7.2.1.2.2. uint32_t hdr_size | Address size - PE ONLY |
| 7.2.1.2.3. uint16_t nsections | Number of sections |
| 7.2.1.2.4. uint32_t offset
(nonzero if embedded) | Offset where this executable start in file |
| 7.2.1.2.5. uint32_t res_addr | Resources RVA - PE ONLY |
| 7.2.1.2.6. struct cli_exe_section* section | Information about all the sections of this file. This array has nsection elements |
| 7.2.1.2.7. struct cli_hashset* vinfo | Hashset for versioninfo matching |

7.2.2. cli_exe_section Struct Reference

Data Fields

- uint32_t [rva](#)
- uint32_t [vsz](#)
- uint32_t [raw](#)
- uint32_t [rsz](#)
- uint32_t [chr](#)
- uint32_t [urva](#)
- uint32_t [uvsz](#)
- uint32_t [uraw](#)
- uint32_t [ursz](#)

7.2.2.1. Detailed Description

Section of executable file.

[PE](#)

7.2.2.2. Field Documentation

7.2.2.2.1. <code>uint32_t chr</code>	Section characteristics
7.2.2.2.2. <code>uint32_t raw</code>	Raw offset (in file)
7.2.2.2.3. <code>uint32_t rsz</code>	Raw size (in file)
7.2.2.2.4. <code>uint32_t rva</code>	Relative VirtualAddress
7.2.2.2.5. <code>uint32_t uraw</code>	PE - unaligned PointerToRawData
7.2.2.2.6. <code>uint32_t ursz</code>	PE - unaligned SizeOfRawData
7.2.2.2.7. <code>uint32_t urva</code>	PE - unaligned VirtualAddress
7.2.2.2.8. <code>uint32_t uvsz</code>	PE - unaligned VirtualSize
7.2.2.2.9. <code>uint32_t vsz</code>	VirtualSize

7.2.3. `cli_pe_hook_data` Struct Reference

Data Fields

- `uint32_t ep`
- `uint16_t nsections`
- struct `pe_image_file_hdr file_hdr`
- struct `pe_image_optional_hdr32 opt32`
- struct `pe_image_optional_hdr64 opt64`
- struct `pe_image_data_dir dirs` [16]
- `uint32_t e_lfanew`
- `uint32_t overlays`
- `int32_t overlays_sz`
- `uint32_t hdr_size`

7.2.3.1. Detailed Description

Data for the bytecode PE hook

[PE](#)

7.2.3.2. Field Documentation

- 7.2.3.2.1. struct pe_image_data_dir dirs[16]** PE data directory header
- 7.2.3.2.2. uint32_t e_lfanew** address of new exe header
- 7.2.3.2.3. uint32_t ep** EntryPoint as file offset
- 7.2.3.2.4. struct pe_image_file_hdr file_hdr** Header for this PE file
- 7.2.3.2.5. uint32_t hdr_size** internally needed by rawaddr
- 7.2.3.2.6. uint16_t nsections** Number of sections
- 7.2.3.2.7. struct pe_image_optional_hdr32 opt32** 32-bit PE optional header
- 7.2.3.2.8. struct pe_image_optional_hdr64 opt64** 64-bit PE optional header
- 7.2.3.2.9. uint32_t overlays** number of overlays
- 7.2.3.2.10. int32_t overlays_sz** size of overlays

7.2.4. DIS_arg Struct Reference

Data Fields

- enum [DIS_ACCESS](#) access_type
- enum [DIS_SIZE](#) access_size
- struct [DIS_mem_arg](#) mem
- enum [X86REGS](#) reg
- uint64_t other

7.2.4.1. Detailed Description

disassembled operand

[Disassemble](#)

7.2.4.2. Field Documentation

- | | |
|--|------------------|
| 7.2.4.2.1. enum DIS_SIZE access_size | size of access |
| 7.2.4.2.2. enum DIS_ACCESS access_type | type of access |
| 7.2.4.2.3. struct DIS_mem_arg mem | memory operand |
| 7.2.4.2.4. uint64_t other | other operand |
| 7.2.4.2.5. enum X86REGS reg | register operand |

7.2.5. DIS_fixed Struct Reference

Data Fields

- enum [X86OPS](#) x86_opcode
- enum [DIS_SIZE](#) operation_size
- enum [DIS_SIZE](#) address_size
- uint8_t [segment](#)

7.2.5.1. Detailed Description

disassembled instruction.

[Disassemble](#)

7.2.5.2. Field Documentation

- | | |
|---|-------------------|
| 7.2.5.2.1. enum DIS_SIZE address_size | size of address |
| 7.2.5.2.2. enum DIS_SIZE operation_size | size of operation |
| 7.2.5.2.3. uint8_t segment | segment |

7.2.5.2.4. enum X86OPS x86_opcode opcode of X86 instruction

7.2.6. DIS_mem_arg Struct Reference

Data Fields

- enum [DIS_SIZE](#) `access_size`
- enum [X86REGS](#) `scale_reg`
- enum [X86REGS](#) `add_reg`
- uint8_t `scale`
- int32_t `displacement`

7.2.6.1. Detailed Description

disassembled memory operand: `scale_reg*scale + add_reg + displacement`

[Disassemble](#)

7.2.6.2. Field Documentation

7.2.6.2.1. enum DIS_SIZE access_size size of access

7.2.6.2.2. enum X86REGS add_reg register used as displacement

7.2.6.2.3. int32_t displacement displacement as immediate number

7.2.6.2.4. uint8_t scale scale as immediate number

7.2.6.2.5. enum X86REGS scale_reg register used as scale

7.2.7. DISASM_RESULT Struct Reference

7.2.7.1. Detailed Description

disassembly result, 64-byte, matched by type-8 signatures

7.2.8. pe_image_data_dir Struct Reference

7.2.8.1. Detailed Description

PE data directory header

[PE](#)

7.2.9. `pe_image_file_hdr` Struct Reference

Data Fields

- `uint32_t` [Magic](#)
- `uint16_t` [Machine](#)
- `uint16_t` [NumberOfSections](#)
- `uint32_t` [TimeDateStamp](#)
- `uint32_t` [PointerToSymbolTable](#)
- `uint32_t` [NumberOfSymbols](#)
- `uint16_t` [SizeOfOptionalHeader](#)

7.2.9.1. Detailed Description

Header for this PE file

[PE](#)

7.2.9.2. Field Documentation

- | | |
|--|---|
| 7.2.9.2.1. <code>uint16_t</code> Machine | CPU this executable runs on, see libclamav/pe.c for possible values |
| 7.2.9.2.2. <code>uint32_t</code> Magic | PE magic header: <code>PE\0\0</code> |
| 7.2.9.2.3. <code>uint16_t</code> NumberOfSections | Number of sections in this executable |
| 7.2.9.2.4. <code>uint32_t</code> NumberOfSymbols | debug |
| 7.2.9.2.5. <code>uint32_t</code> PointerToSymbolTable | debug |
| 7.2.9.2.6. <code>uint16_t</code> SizeOfOptionalHeader | == 224 |
| 7.2.9.2.7. <code>uint32_t</code> TimeDateStamp | Unreliable |

7.2.10. `pe_image_optional_hdr32` Struct Reference

Data Fields

- `uint8_t` [MajorLinkerVersion](#)
- `uint8_t` [MinorLinkerVersion](#)
- `uint32_t` [SizeOfCode](#)
- `uint32_t` [SizeOfInitializedData](#)
- `uint32_t` [SizeOfUninitializedData](#)
- `uint32_t` [ImageBase](#)
- `uint32_t` [SectionAlignment](#)
- `uint32_t` [FileAlignment](#)
- `uint16_t` [MajorOperatingSystemVersion](#)

- uint16_t [MinorOperatingSystemVersion](#)
- uint16_t [MajorImageVersion](#)
- uint16_t [MinorImageVersion](#)
- uint32_t [Checksum](#)
- uint32_t [NumberOfRvaAndSizes](#)

7.2.10.1. Detailed Description

32-bit PE optional header

[PE](#)

7.2.10.2. Field Documentation

7.2.10.2.1. uint32_t CheckSum	NT drivers only
7.2.10.2.2. uint32_t FileAlignment	usually 32 or 512
7.2.10.2.3. uint32_t ImageBase	multiple of 64 KB
7.2.10.2.4. uint16_t MajorImageVersion	unreliable
7.2.10.2.5. uint8_t MajorLinkerVersion	unreliable
7.2.10.2.6. uint16_t MajorOperatingSystemVersion	not used
7.2.10.2.7. uint16_t MinorImageVersion	unreliable
7.2.10.2.8. uint8_t MinorLinkerVersion	unreliable
7.2.10.2.9. uint16_t MinorOperatingSystemVersion	not used
7.2.10.2.10. uint32_t NumberOfRvaAndSizes	unreliable
7.2.10.2.11. uint32_t SectionAlignment	usually 32 or 4096
7.2.10.2.12. uint32_t SizeOfCode	unreliable
7.2.10.2.13. uint32_t SizeOfInitializedData	unreliable

7.2.10.2.14. uint32_t SizeOfUninitializedData unreliable

7.2.11. pe_image_optional_hdr64 Struct Reference

Data Fields

- uint8_t [MajorLinkerVersion](#)
- uint8_t [MinorLinkerVersion](#)
- uint32_t [SizeOfCode](#)
- uint32_t [SizeOfInitializedData](#)
- uint32_t [SizeOfUninitializedData](#)
- uint64_t [ImageBase](#)
- uint32_t [SectionAlignment](#)
- uint32_t [FileAlignment](#)
- uint16_t [MajorOperatingSystemVersion](#)
- uint16_t [MinorOperatingSystemVersion](#)
- uint16_t [MajorImageVersion](#)
- uint16_t [MinorImageVersion](#)
- uint32_t [Checksum](#)
- uint32_t [NumberOfRvaAndSizes](#)

7.2.11.1. Detailed Description

PE 64-bit optional header

[PE](#)

7.2.11.2. Field Documentation

- | | |
|---|-------------------|
| 7.2.11.2.1. uint32_t CheckSum | NT drivers only |
| 7.2.11.2.2. uint32_t FileAlignment | usually 32 or 512 |
| 7.2.11.2.3. uint64_t ImageBase | multiple of 64 KB |
| 7.2.11.2.4. uint16_t MajorImageVersion | unreliable |
| 7.2.11.2.5. uint8_t MajorLinkerVersion | unreliable |
| 7.2.11.2.6. uint16_t MajorOperatingSystemVersion | not used |

7.2.11.2.7. uint16_t MinorImageVersion	unreliable
7.2.11.2.8. uint8_t MinorLinkerVersion	unreliable
7.2.11.2.9. uint16_t MinorOperatingSystemVersion	not used
7.2.11.2.10. uint32_t NumberOfRvaAndSizes	unreliable
7.2.11.2.11. uint32_t SectionAlignment	usually 32 or 4096
7.2.11.2.12. uint32_t SizeOfCode	unreliable
7.2.11.2.13. uint32_t SizeOfInitializedData	unreliable
7.2.11.2.14. uint32_t SizeOfUninitializedData	unreliable

7.2.12. `pe_image_section_hdr` Struct Reference

Data Fields

- `uint8_t Name` [8]
- `uint32_t SizeOfRawData`
- `uint32_t PointerToRawData`
- `uint32_t PointerToRelocations`
- `uint32_t PointerToLinenumbers`
- `uint16_t NumberOfRelocations`
- `uint16_t NumberOfLinenumbers`

7.2.12.1. Detailed Description

PE section header

PE

7.2.12.2. Field Documentation

7.2.12.2.1. uint8_t Name[8]	may not end with NULL
7.2.12.2.2. uint16_t NumberOfLinenumbers	object files only
7.2.12.2.3. uint16_t NumberOfRelocations	object files only

7.2.12.2.4. <code>uint32_t</code> <code>PointerToLinenumbers</code>	object files only
7.2.12.2.5. <code>uint32_t</code> <code>PointerToRawData</code>	offset to the section's data
7.2.12.2.6. <code>uint32_t</code> <code>PointerToRelocations</code>	object files only
7.2.12.2.7. <code>uint32_t</code> <code>SizeOfRawData</code>	multiple of <code>FileAlignment</code>

7.3. Low level API

7.3.1. `bytecode_api.h` File Reference

Enumerations

- enum `BytecodeKind` { `BC_GENERIC` = 0 , `BC_LOGICAL` = 256, `BC_PE_UNPACKER` }
- enum { `PE_INVALID_RVA` = 0xFFFFFFFF }
- enum `FunctionalityLevels`
- enum `pdf_phase`
- enum `pdf_flag`
- enum `pdf_objflags`
- enum { `SEEK_SET` = 0, `SEEK_CUR`, `SEEK_END` }

Functions

- `uint32_t test1` (`uint32_t` a, `uint32_t` b)
- `int32_t read` (`uint8_t` *data, `int32_t` size)
Reads specified amount of bytes from the current file into a buffer. Also moves current position in the file.
- `int32_t write` (`uint8_t` *data, `int32_t` size)
Writes the specified amount of bytes from a buffer to the current temporary file.
- `int32_t seek` (`int32_t` pos, `uint32_t` whence)
Changes the current file position to the specified one.
- `uint32_t setvirusname` (const `uint8_t` *name, `uint32_t` len)
- `uint32_t debug_print_str` (const `uint8_t` *str, `uint32_t` len)
- `uint32_t debug_print_uint` (`uint32_t` a)
- `uint32_t disasm_x86` (struct `DISASM_RESULT` *result, `uint32_t` len)
- `uint32_t pe_rawaddr` (`uint32_t` rva)
- `int32_t file_find` (const `uint8_t` *data, `uint32_t` len)
- `int32_t file_byteat` (`uint32_t` offset)
- void * `malloc` (`uint32_t` size)
- `uint32_t test2` (`uint32_t` a)

- `int32_t get_pe_section` (`struct cli_exe_section *section, uint32_t num`)
- `int32_t fill_buffer` (`uint8_t *buffer, uint32_t len, uint32_t filled, uint32_t cursor, uint32_t fill`)
- `int32_t extract_new` (`int32_t id`)
- `int32_t read_number` (`uint32_t radix`)
- `int32_t hashset_new` (`void`)
- `int32_t hashset_add` (`int32_t hs, uint32_t key`)
- `int32_t hashset_remove` (`int32_t hs, uint32_t key`)
- `int32_t hashset_contains` (`int32_t hs, uint32_t key`)
- `int32_t hashset_done` (`int32_t id`)
- `int32_t hashset_empty` (`int32_t id`)
- `int32_t buffer_pipe_new` (`uint32_t size`)
- `int32_t buffer_pipe_new_fromfile` (`uint32_t pos`)
- `uint32_t buffer_pipe_read_avail` (`int32_t id`)
- `uint8_t * buffer_pipe_read_get` (`int32_t id, uint32_t amount`)
- `int32_t buffer_pipe_read_stopped` (`int32_t id, uint32_t amount`)
- `uint32_t buffer_pipe_write_avail` (`int32_t id`)
- `uint8_t * buffer_pipe_write_get` (`int32_t id, uint32_t size`)
- `int32_t buffer_pipe_write_stopped` (`int32_t id, uint32_t amount`)
- `int32_t buffer_pipe_done` (`int32_t id`)
- `int32_t inflate_init` (`int32_t from_buffer, int32_t to_buffer, int32_t windowBits`)
- `int32_t inflate_process` (`int32_t id`)
- `int32_t inflate_done` (`int32_t id`)
- `int32_t bytecode_rt_error` (`int32_t locationid`)
- `int32_t jsnorm_init` (`int32_t from_buffer`)
- `int32_t jsnorm_process` (`int32_t id`)
- `int32_t jsnorm_done` (`int32_t id`)
- `int32_t ilog2` (`uint32_t a, uint32_t b`)
- `int32_t ipow` (`int32_t a, int32_t b, int32_t c`)
- `uint32_t iexp` (`int32_t a, int32_t b, int32_t c`)
- `int32_t isin` (`int32_t a, int32_t b, int32_t c`)
- `int32_t icos` (`int32_t a, int32_t b, int32_t c`)
- `int32_t memstr` (`const uint8_t *haystack, int32_t haysize, const uint8_t *needle, int32_t needlesize`)
- `int32_t hex2ui` (`uint32_t hex1, uint32_t hex2`)
- `int32_t atoi` (`const uint8_t *str, int32_t size`)
- `uint32_t debug_print_str_start` (`const uint8_t *str, uint32_t len`)
- `uint32_t debug_print_str_nonl` (`const uint8_t *str, uint32_t len`)
- `uint32_t entropy_buffer` (`uint8_t *buffer, int32_t size`)
- `int32_t map_new` (`int32_t keysize, int32_t valuesize`)
- `int32_t map_addkey` (`const uint8_t *key, int32_t ksize, int32_t id`)
- `int32_t map_setvalue` (`const uint8_t *value, int32_t vsize, int32_t id`)
- `int32_t map_remove` (`const uint8_t *key, int32_t ksize, int32_t id`)
- `int32_t map_find` (`const uint8_t *key, int32_t ksize, int32_t id`)
- `int32_t map_getvaluesize` (`int32_t id`)
- `uint8_t * map_getvalue` (`int32_t id, int32_t size`)

- `int32_t map_done` (`int32_t id`)
- `int32_t file_find_limit` (`const uint8_t *data`, `uint32_t len`, `int32_t max-pos`)
- `uint32_t engine_functionality_level` (`void`)
- `uint32_t engine_dconf_level` (`void`)
- `uint32_t engine_scan_options` (`void`)
- `uint32_t engine_db_options` (`void`)
- `int32_t extract_set_container` (`uint32_t container`)
- `int32_t input_switch` (`int32_t extracted_file`)
- `uint32_t get_environment` (`struct cli_environment *env`, `uint32_t len`)
- `uint32_t disable_bytecode_if` (`const int8_t *reason`, `uint32_t len`, `uint32_t cond`)
- `uint32_t disable_jit_if` (`const int8_t *reason`, `uint32_t len`, `uint32_t cond`)
- `int32_t version_compare` (`const uint8_t *lhs`, `uint32_t lhs_len`, `const uint8_t *rhs`, `uint32_t rhs_len`)
- `uint32_t check_platform` (`uint32_t a`, `uint32_t b`, `uint32_t c`)
- `int32_t pdf_get_obj_num` (`void`)
- `int32_t pdf_get_flags` (`void`)
- `int32_t pdf_set_flags` (`int32_t flags`)
- `int32_t pdf_lookupobj` (`uint32_t id`)
- `uint32_t pdf_getobjsize` (`int32_t objidx`)
- `uint8_t * pdf_getobj` (`int32_t objidx`, `uint32_t amount`)
- `int32_t pdf_get_phase` (`void`)
- `int32_t pdf_get_dumpedobjid` (`void`)
- `int32_t matchicon` (`const uint8_t *group1`, `int32_t group1_len`, `const uint8_t *group2`, `int32_t group2_len`)

Variables

- `const uint32_t __clambc_match_counts` [64]
Logical signature match counts.
- `const uint32_t __clambc_match_offsets` [64]
Logical signature match offsets This is a low-level variable, use the Macros in `bytecode_local.h` instead to access it.
- `struct cli_pe_hook_data __clambc_pdata`
- `const uint32_t __clambc_filesize` [1]
- `const uint16_t __clambc_kind`

7.3.1.1. Detailed Description

7.3.1.2. Enumeration Type Documentation

7.3.1.2.1. anonymous enum

Enumerator:

`PE_INVALID_RVA` Invalid RVA specified

7.3.1.2.2. anonymous enum**Enumerator:**

SEEK_SET set file position to specified absolute position
SEEK_CUR set file position relative to current position
SEEK_END set file position relative to file end

7.3.1.2.3. enum BytecodeKind Bytecode trigger kind**Enumerator:**

BC_GENERIC generic bytecode, not tied a specific hook
BC_LOGICAL triggered by a logical signature
BC_PE_UNPACKER a PE unpacker

7.3.1.2.4. enum FunctionalityLevels LibClamAV functionality level constants**7.3.1.2.5. enum pdf_flag** PDF flags**7.3.1.2.6. enum pdf_objflags** PDF obj flags**7.3.1.2.7. enum pdf_phase** Phase of PDF parsing**7.3.1.3. Function Documentation**

7.3.1.3.1. int32_t atoi (const uint8_t * *str*, int32_t *size*) Converts string to positive number.

Parameters

str buffer
size size of *str*

Returns

>0 string converted to number if possible, -1 on error

String operation

7.3.1.3.2. int32_t buffer_pipe_done (int32_t *id*) Deallocate memory used by buffer.

Data structure

After this all attempts to use this buffer will result in error. All *buffer_pipes* are automatically deallocated when bytecode finishes execution.

Parameters

id ID of *buffer_pipe*

Returns

0 on success

7.3.1.3.3. `int32_t buffer_pipe_new (uint32_t size)` Creates a new pipe with the specified buffer size

Data structure

Parameters

size size of buffer

Returns

ID of newly created `buffer_pipe`

7.3.1.3.4. `int32_t buffer_pipe_new_fromfile (uint32_t pos)` Same as `buffer_pipe_new`, except the pipe's input is tied

Data structure

File operation

to the current file, at the specified position.

Parameters

pos starting position of pipe input in current file

Returns

ID of newly created `buffer_pipe`

7.3.1.3.5. `uint32_t buffer_pipe_read_avail (int32_t id)` Returns the amount of bytes available to read.

Data structure

Parameters

id ID of `buffer_pipe`

Returns

amount of bytes available to read

7.3.1.3.6. `uint8_t* buffer_pipe_read_get (int32_t id, uint32_t amount)` Returns a pointer to the buffer for reading.

Data structure

The 'amount' parameter should be obtained by a call to `buffer_pipe_read_avail()`.

Parameters

id ID of `buffer_pipe`

amount to read

Returns

pointer to buffer, or NULL if buffer has less than specified amount

7.3.1.3.7. `int32_t buffer_pipe_read_stopped (int32_t id, uint32_t amount)`

Data structure

Updates read cursor in `buffer_pipe`.

Parameters

id ID of `buffer_pipe`

amount amount of bytes to move read cursor

Returns

0 on success

7.3.1.3.8. `uint32_t buffer_pipe_write_avail (int32_t id)` Returns the amount of bytes available for writing.

Data structure

Parameters

id ID of `buffer_pipe`

Returns

amount of bytes available for writing

7.3.1.3.9. `uint8_t* buffer_pipe_write_get (int32_t id, uint32_t size)`

Data structure

Returns pointer to writable buffer. The 'amount' parameter should be obtained by a call to [buffer_pipe_write_avail\(\)](#).

Parameters

id ID of `buffer_pipe`

size amount of bytes to write

Returns

pointer to write buffer, or NULL if requested amount is more than what is available in the buffer

7.3.1.3.10. `int32_t buffer_pipe_write_stopped (int32_t id, uint32_t amount)` Updates the write cursor in `buffer_pipe`.

Data structure

Parameters

id ID of `buffer_pipe`

amount amount of bytes to move write cursor

Returns

0 on success

7.3.1.3.11. `int32_t` `bytecode_rt_error` (`int32_t` *locationid*) Report a runtime error at the specified locationID.

Scan

Parameters

locationid (line << 8) | (column&0xff)

Returns

0

7.3.1.3.12. `uint32_t` `check_platform` (`uint32_t` *a*, `uint32_t` *b*, `uint32_t` *c*) Disables the JIT if the platform id matches. 0xff can be used instead of a field to mark ANY.

Parameters

a - os_category << 24 | arch << 20 | compiler << 16 | flevel << 8 | dconf

b - big_endian << 28 | sizeof_ptr << 24 | cpp_version

c - os_features << 24 | c_version

Returns

0 - no match 1 - match

Environment

7.3.1.3.13. `uint32_t` `debug_print_str` (`const uint8_t *` *str*, `uint32_t` *len*) Prints a debug message.

Parameters

[in] *str* Message to print

[in] *len* length of message to print

Returns

0

String operation

7.3.1.3.14. `uint32_t` `debug_print_str_nonl` (`const uint8_t *` *str*, `uint32_t` *len*) Prints a debug message with a trailing newline, and not preceded by 'LibClamAV debug'.

Parameters

str the string

len length of *str*

Returns

0

String operation

7.3.1.3.15. `uint32_t debug_print_str_start (const uint8_t * str, uint32_t len)` Prints a debug message with a trailing newline, but preceded by 'LibClamAV debug'.

Parameters*str* the string*len* length of *str***Returns**

0

String operation

7.3.1.3.16. `uint32_t debug_print_uint (uint32_t a)` Prints a number as a debug message. This is like `debug_print_str_nonl!`

Parameters[in] *a* number to print**Returns**

0

String operation

7.3.1.3.17. `uint32_t disable_bytecode_if (const int8_t * reason, uint32_t len, uint32_t cond)` Disables the bytecode completely if condition is true. Can only be called from the BC_STARTUP bytecode.

Parameters*reason* - why the bytecode had to be disabled*len* - length of reason*cond* - condition**Returns**

0 - auto mode 1 - JIT disabled 2 - fully disabled

Environment

7.3.1.3.18. `uint32_t disable_jit_if (const int8_t * reason, uint32_t len, uint32_t cond)` Disables the JIT completely if condition is true. Can only be called from the BC_STARTUP bytecode.

Parameters

reason - why the JIT had to be disabled

len - length of reason

cond - condition

Returns

0 - auto mode 1 - JIT disabled 2 - fully disabled

Environment

7.3.1.3.19. `uint32_t disasm_x86 (struct DISASM_RESULT * result, uint32_t len)` Disassembles starting from current file position, the specified amount of bytes.

Parameters

[out] *result* pointer to struct holding result

[in] *len* how many bytes to disassemble

Returns

0 for success

You can use `lseek` to disassemble starting from a different location. This is a low-level API, the result is in ClamAV type-8 signature format (64 bytes/instruction).

See also

[DisassembleAt](#)

Disassemble

7.3.1.3.20. `uint32_t engine_db_options (void)` Returns the current engine's db options.

Returns

CL_DB_* flags

Engine query

7.3.1.3.21. `uint32_t engine_dconf_level (void)` Returns the current engine (dconf) functionality level. Usually identical to [engine_functionality_level\(\)](#), unless distro backported patches. Compare with [FunctionalityLevels](#).

Returns

an integer representing the DCONF (security fixes) level.

Engine query

7.3.1.3.22. `uint32_t engine_functionality_level (void)` Returns the current engine (feature) functionality level. To map these to ClamAV releases, compare it with [FunctionalityLevels](#).

Returns

an integer representing current engine functionality level.

Engine query

7.3.1.3.23. `uint32_t engine_scan_options (void)` Returns the current engine's scan options.

Returns

CL_SCAN* flags

Engine query

7.3.1.3.24. `uint32_t entropy_buffer (uint8_t * buffer, int32_t size)` Returns an approximation for the entropy of *buffer*.

Parameters

buffer input buffer

size size of buffer

Returns

entropy estimation * 2²⁶

String operation

7.3.1.3.25. `int32_t extract_new (int32_t id)` Prepares for extracting a new file, if we've already extracted one it scans it.

Scan

Parameters

[in] *id* an id for the new file (for example position in container)

Returns

1 if previous extracted file was infected

7.3.1.3.26. int32_t extract_set_container (uint32_t *container*)

Sets the container type for the currently extracted file.

Parameters

container container type (CL_TYPE_*)

Returns

current setting for container (CL_TYPE_ANY default)

Scan**7.3.1.3.27. int32_t file_byteat (uint32_t *offset*)** Read a single byte from current file**File operation****Parameters**

offset file offset

Returns

byte at offset *off* in the current file, or -1 if offset is invalid

7.3.1.3.28. int32_t file_find (const uint8_t * *data*, uint32_t *len*) Looks for the specified sequence of bytes in the current file.**File operation****Parameters**

[in] *data* the sequence of bytes to look for

len length of *data*, cannot be more than 1024

Returns

offset in the current file if match is found, -1 otherwise

7.3.1.3.29. `int32_t file_find_limit (const uint8_t * data, uint32_t len, int32_t maxpos)` Looks for the specified sequence of bytes in the current file, up to the specified position.

Parameters

- [in] *data* the sequence of bytes to look for
- len* length of *data*, cannot be more than 1024
- maxpos* maximum position to look for a match, note that this is 1 byte after the end of last possible match: $\text{match_pos} + \text{len} < \text{maxpos}$

Returns

offset in the current file if match is found, -1 otherwise

File operation

7.3.1.3.30. `int32_t fill_buffer (uint8_t * buffer, uint32_t len, uint32_t filled, uint32_t cursor, uint32_t fill)` Fills the specified buffer with at least *fill* bytes.

File operation

Parameters

- [out] *buffer* the buffer to fill
- [in] *len* length of buffer
- [in] *filled* how much of the buffer is currently filled
- [in] *cursor* position of cursor in buffer
- [in] *fill* amount of bytes to fill in (0 is valid)

Returns

<0 on error, 0 on EOF, number bytes available in buffer (starting from 0)
The character at the cursor will be at position 0 after this call.

7.3.1.3.31. `uint32_t get_environment (struct cli_environment * env, uint32_t len)` Queries the environment this bytecode runs in. Used by BC_STARTUP to disable bytecode when bugs are known for the current platform.

Parameters

- [out] *env* - the full environment
- len* - size of *env*

Returns

0

Environment

7.3.1.3.32. `int32_t get_pe_section (struct cli_exe_section * section, uint32_t num)` Gets information about the specified PE section.

PE

Parameters

[out] *section* PE section information will be stored here

[in] *num* PE section number

Returns

0 - success -1 - failure

7.3.1.3.33. `int32_t hashset_add (int32_t hs, uint32_t key)` Add a new 32-bit key to the hashset.

Data structure

Parameters

hs ID of hashset (from `hashset_new`)

key the key to add

Returns

0 on success

7.3.1.3.34. `int32_t hashset_contains (int32_t hs, uint32_t key)` Returns whether the hashset contains the specified key.

Data structure

Parameters

hs ID of hashset (from `hashset_new`)

key the key to lookup

Returns

1 if found, 0 if not found, <0 on invalid hashset ID

7.3.1.3.35. `int32_t hashset_done (int32_t id)` Deallocates the memory used by the specified hashset.

Data structure

Trying to use the hashset after this will result in an error. The hashset may not be used after this. All hashsets are automatically deallocated when bytecode finishes execution.

Parameters

id ID of hashset (from `hashset_new`)

Returns

0 on success

7.3.1.3.36. `int32_t` `hashset__empty` (`int32_t` *id*) Returns whether the hashset is empty.

Data structure

Parameters

id of hashset (from `hashset__new`)

Returns

0 on success

7.3.1.3.37. `int32_t` `hashset__new` (`void`) Creates a new hashset and returns its id.

Data structure

Returns

ID for new hashset

7.3.1.3.38. `int32_t` `hashset__remove` (`int32_t` *hs*, `uint32_t` *key*) Remove a 32-bit key from the hashset.

Data structure

Parameters

hs ID of hashset (from `hashset__new`)

key the key to add

Returns

0 on success

7.3.1.3.39. `int32_t` `hex2ui` (`uint32_t` *hex1*, `uint32_t` *hex2*) Returns hexadecimal characters *hex1* and *hex2* converted to 8-bit number.

Parameters

hex1 hexadecimal character

hex2 hexadecimal character

Returns

hex1 hex2 converted to 8-bit integer, -1 on error

String operation

7.3.1.3.40. int32_t icos (int32_t a, int32_t b, int32_t c)Returns $c \cdot \cos(a/b)$.**Parameters***a* integer*b* integer*c* integer**Returns** $c \cdot \sin(a/b)$ **Math function****7.3.1.3.41. uint32_t iexp (int32_t a, int32_t b, int32_t c)**Returns $\exp(a/b) \cdot c$ **Parameters***a* integer*b* integer*c* integer**Returns** $c \cdot \exp(a/b)$ **Math function****7.3.1.3.42. int32_t ilog2 (uint32_t a, uint32_t b)** Returns $2^{26} \cdot \log_2(a/b)$ **Parameters***a* input*b* input**Returns** $2^{26} \cdot \log_2(a/b)$ **Math function**

7.3.1.3.43. int32_t inflate_done (int32_t id) Deallocates inflate data structure. Using the inflate data structure after this will result in an error. All inflate data structures are automatically deallocated when bytecode finishes execution.

Data structure

Parameters

id ID of inflate data structure

Returns

0 on success.

7.3.1.3.44. `int32_t inflate_init (int32_t from_buffer, int32_t to_buffer, int32_t windowBits)` Initializes inflate data structures for decompressing data

Data structure

'from_buffer' and writing uncompressed data 'to_buffer'.

Parameters

from_buffer ID of buffer_pipe to read compressed data from

to_buffer ID of buffer_pipe to write decompressed data to

windowBits (see zlib documentation)

Returns

ID of newly created inflate data structure, <0 on failure

7.3.1.3.45. `int32_t inflate_process (int32_t id)` Inflate all available data in the input buffer, and write to output buffer. Stops when the input buffer becomes empty, or write buffer becomes full. Also attempts to recover from corrupted inflate stream (via `inflateSync`). This function can be called repeatedly on success after filling the input buffer, and flushing the output buffer. The inflate stream is done processing when 0 bytes are available from output buffer, and input buffer is not empty.

Data structure**Parameters**

id ID of inflate data structure

Returns

0 on success, zlib error code otherwise

7.3.1.3.46. `int32_t input_switch (int32_t extracted_file)` Toggles the read/seek API to read from the currently extracted file, and back. You must call seek after switching inputs to position the cursor to a valid position.

Parameters

extracted_file 1 - switch to reading from extracted file, 0 - switch back to original input

Returns

-1 on error (if no extracted file exists) 0 on success

Scan

7.3.1.3.47. `int32_t ipow (int32_t a, int32_t b, int32_t c)`Returns $c*a^b$.**Parameters***a* integer*b* integer*c* integer**Returns** $c*\text{pow}(a,b)$ **Math function****7.3.1.3.48. `int32_t isin (int32_t a, int32_t b, int32_t c)`**Returns $c*\sin(a/b)$.**Parameters***a* integer*b* integer*c* integer**Returns** $c*\sin(a/b)$ **Math function****7.3.1.3.49. `int32_t jsnorm_done (int32_t id)`**

Flushes JS

normalizer.

JavaScript**Parameters***id* ID of js normalizer to flush**Returns**

0 - success -1 - failure

7.3.1.3.50. `int32_t jsnorm_init (int32_t from_buffer)` Initializes JS normalizer for reading 'from_buffer'. Normalized JS will be written to a single tempfile, one normalized JS per line, and automatically scanned when the bytecode finishes execution.

JavaScript

Parameters

from_buffer ID of *buffer_pipe* to read javascript from

Returns

ID of JS normalizer, <0 on failure

7.3.1.3.51. `int32_t jsnorm_process (int32_t id)` Normalize all javascript from the input buffer, and write to tempfile. You can call this function repeatedly on success, if you (re)fill the input buffer.

JavaScript**Parameters**

id ID of JS normalizer

Returns

0 on success, <0 on failure

7.3.1.3.52. `void* malloc (uint32_t size)` Allocates memory. Currently this memory is freed automatically on exit from the bytecode, and there is no way to free it sooner.

Data structure**Parameters**

size amount of memory to allocate in bytes

Returns

pointer to allocated memory

7.3.1.3.53. `int32_t map_addkey (const uint8_t * key, int32_t ksize, int32_t id)` Inserts the specified key/value pair into the map.

Parameters

id id of table

key key

ksize size of key

Returns

0 - if key existed before 1 - if key didn't exist before <0 - if ksize doesn't match keysize specified at table creation

Data structure

7.3.1.3.54. `int32_t map_done (int32_t id)` Deallocates the memory used by the specified map. Trying to use the map after this will result in an error. All maps are automatically deallocated when the bytecode finishes execution.

Parameters

id id of map

Returns

0 - success -1 - invalid map

Data structure

7.3.1.3.55. `int32_t map_find (const uint8_t * key, int32_t ksize, int32_t id)` Looks up key in map. The map remember the last looked up key (so you can retrieve the value).

Parameters

id id of map

key key

ksize size of key

Returns

0 - if not found 1 - if found <0 - if ksize doesn't match the size specified at table creation

Data structure

7.3.1.3.56. `uint8_t* map_getvalue (int32_t id, int32_t size)` Returns the value obtained during last `map_find`.

Parameters

id id of map.

size size of value (obtained from `map_getvaluesize`)

Returns

value

Data structure

7.3.1.3.57. `int32_t map_getvaluesize (int32_t id)` Returns the size of value obtained during last `map_find`.

Parameters

id id of map.

Returns

size of value

Data structure

7.3.1.3.58. `int32_t map_new (int32_t keysize, int32_t valuesize)` Creates a new map and returns its id.

Parameters

keysize size of key

valuesize size of value, if 0 then value is allocated separately

Returns

ID of new map

Data structure

7.3.1.3.59. `int32_t map_remove (const uint8_t * key, int32_t ksize, int32_t id)` Remove an element from the map.

Parameters

id id of map

key key

ksize size of key

Returns

0 on success, key was present 1 if key was not present <0 if ksize doesn't match keysize specified at table creation

Data structure

7.3.1.3.60. `int32_t map_setvalue (const uint8_t * value, int32_t vsize, int32_t id)` Sets the value for the last inserted key with `map_addkey`.

Parameters

id id of table

value value

vsize size of value

Returns

0 - if update was successful <0 - if there is no last key

Data structure

7.3.1.3.61. `int32_t matchicon (const uint8_t * group1, int32_t group1_len, const uint8_t * group2, int32_t group2_len)` Attempts to match current executable's icon against the specified icon groups.

Icon

Parameters

[in] *group1* - same as GROUP1 in LDB signatures
group1_len - length of *group1*
 [in] *group2* - same as GROUP2 in LDB signatures
group2_len - length of *group2*

Returns

-1 - invalid call, or sizes (only valid for PE hooks) 0 - not a match 1 - match

7.3.1.3.62. `int32_t memstr (const uint8_t * haystack, int32_t haysize, const uint8_t * needle, int32_t needlesize)` Return position of match, -1 otherwise.

Parameters

haystack buffer to search
haysize size of haystack
needle substring to search
needlesize size of needle

Returns

location of match, -1 otherwise

String operation

7.3.1.3.63. `int32_t pdf_get_dumpedobjid (void)` Return the currently dumped obj index.

PDF

Valid only in PDF_PHASE_POSTDUMP.

Returns

>=0 - object index -1 - invalid phase

7.3.1.3.64. `int32_t pdf_get_flags (void)` Return the flags for the entire PDF (as set so far).

Returns

-1 - if not called from PDF hook >=0 - pdf flags

PDF

7.3.1.3.65. `int32_t pdf_get_obj_num (void)` Return number of pdf objects

Returns

-1 - if not called from PDF hook ≥ 0 - number of PDF objects

PDF

7.3.1.3.66. `int32_t pdf_get_phase (void)` Return an 'enum pdf_phase'.

PDF

Identifies at which phase this bytecode was called.

Returns

the current `pdf_phase`

7.3.1.3.67. `uint8_t* pdf_getobj (int32_t objidx, uint32_t amount)` Return the undecoded object.

PDF

Meant only for reading, write modifies the fmap buffer, so avoid!

Parameters

objidx - object index (from 0), not object id!

amount - size returned by `pdf_getobjsize` (or smaller)

Returns

NULL - invalid *objidx/amount* pointer - pointer to original object

7.3.1.3.68. `uint32_t pdf_getobjsize (int32_t objidx)` Return the size of the specified PDF obj.

PDF

Parameters

objidx - object index (from 0), not object id!

Returns

0 - if not called from PDF hook, or invalid *objnum* ≥ 0 - size of object

7.3.1.3.69. `int32_t pdf_lookupobj (uint32_t id)` Lookup pdf object with specified id.

PDF

Parameters

id - pdf id (objnumber << 8 | generationid)

Returns

-1 - if object id doesn't exist >=0 - object index

7.3.1.3.70. `int32_t pdf_set_flags (int32_t flags)` Sets the flags for the entire PDF. It is recommended that you retrieve old flags, and just add new ones.

PDF**Parameters**

flags - flags to set.

Returns

0 - success -1 - invalid phase

7.3.1.3.71. `uint32_t pe_rawaddr (uint32_t rva)` Converts a RVA (Relative Virtual Address) to an absolute PE file offset.

Parameters

rva a rva address from the PE file

Returns

absolute file offset mapped to the *rva*, or PE_INVALID_RVA if the *rva* is invalid.

PE

7.3.1.3.72. `int32_t read (uint8_t * data, int32_t size)` Reads specified amount of bytes from the current file into a buffer. Also moves current position in the file.

Parameters

[in] *size* amount of bytes to read

[out] *data* pointer to buffer where data is read into

Returns

amount read.

File operation

7.3.1.3.73. `int32_t read_number (uint32_t radix)` Reads a number in the specified radix starting from the current position.

File operation

Non-numeric characters are ignored.

Parameters

[in] *radix* 10 or 16

Returns

the number read

7.3.1.3.74. `int32_t seek (int32_t pos, uint32_t whence)` Changes the current file position to the specified one.

See also

[SEEK_SET](#), [SEEK_CUR](#), [SEEK_END](#)

Parameters

[in] *pos* offset (absolute or relative depending on *whence* param)

[in] *whence* one of `SEEK_SET`, `SEEK_CUR`, `SEEK_END`

Returns

absolute position in file

File operation

7.3.1.3.75. `uint32_t setvirusname (const uint8_t * name, uint32_t len)` Sets the name of the virus found.

Parameters

[in] *name* the name of the virus

[in] *len* length of the virusname

Returns

0

Scan

7.3.1.3.76. `uint32_t test1 (uint32_t a, uint32_t b)` Test api.

Parameters

a 0xf00dbeef

b 0xbeeff00d

Returns

0x12345678 if parameters match, 0x55 otherwise

7.3.1.3.77. uint32_t test2 (uint32_t *a*) Test api2.

Parameters

a 0xf00d

Returns

0xd00f if parameter matches, 0x5555 otherwise

7.3.1.3.78. int32_t version_compare (const uint8_t * *lhs*, uint32_t *lhs_len*, const uint8_t * *rhs*, uint32_t *rhs_len*) Compares two version numbers.

Parameters

[in] *lhs* - left hand side of comparison

lhs_len - length of lhs

[in] *rhs* - right hand side of comparison

rhs_len - length of rhs

Returns

-1 - lhs < rhs 0 - lhs == rhs 1 - lhs > rhs

Environment

7.3.1.3.79. int32_t write (uint8_t * *data*, int32_t *size*)

Writes the specified amount of bytes from a buffer to the current temporary file.

Parameters

[in] *data* pointer to buffer of data to write

[in] *size* amount of bytes to write *size* bytes to temporary file, from the buffer pointed to byte

Returns

amount of bytes successfully written

File operation

7.3.1.4. Variable Documentation

7.3.1.4.1. const uint32_t __clambc__filesize[1] File size (max 4G).

Global variable

7.3.1.4.2. `const uint16_t __clambc_kind` Kind of the bytecode

Global variable

7.3.1.4.3. `const uint32_t __clambc_match_counts[64]`

Logical signature match counts.

This is a low-level variable, use the Macros in [bytecode_local.h](#) instead to access it.

Global variable

7.3.1.4.4. `const uint32_t __clambc_match_offsets[64]`

Logical signature match offsets This is a low-level variable, use the Macros in [bytecode_local.h](#) instead to access it.

Global variable

7.3.1.4.5. `struct cli_pe_hook_data __clambc_pedata` PE data, if this is a PE hook.

Global variable

7.3.2. `bytecode_disasm.h` File Reference

Data Structures

- struct [DISASM_RESULT](#)

Enumerations

- enum [X86OPS](#) { ,
[OP_AAA](#), [OP_AAD](#), [OP_AAM](#), [OP_AAS](#),
[OP_ADD](#), [OP_ADC](#), [OP_AND](#), [OP_ARPL](#),
[OP_BOUND](#), [OP_BSF](#), [OP_BSR](#), [OP_BSWAP](#),
[OP_BT](#), [OP BTC](#), [OP_BTR](#), [OP_BTS](#),
[OP_CALL](#), [OP_CDQ](#) , [OP_CWDE](#), [OP_CBW](#),
[OP_CLC](#), [OP_CLD](#), [OP_CLI](#), [OP_CLTS](#),
[OP_CMC](#), [OP_CMOVO](#), [OP_CMOVNO](#), [OP_CMOVC](#),
[OP_CMOVNC](#), [OP_CMOVZ](#), [OP_CMOVNZ](#), [OP_CMOVBE](#),
[OP_CMOVA](#), [OP_CMOVS](#), [OP_CMOVNS](#), [OP_CMOVP](#),
[OP_CMOVNP](#), [OP_CMOVL](#), [OP_CMOVGE](#), [OP_CMOVLE](#),
[OP_CMOVG](#), [OP_CMP](#), [OP_CMPSD](#), [OP_CMPSW](#),

OP_CMPSB, OP_CMPXCHG, OP_CMPXCHG8B, OP_CPUID,
 OP_DAA, OP_DAS, OP_DEC, OP_DIV,
 OP_ENTER, OP_FWAIT, OP_HLT, OP_IDIV,
 OP_IMUL, OP_INC, OP_IN, OP_INSD,
 OP_INSW, OP_INSB, OP_INT, OP_INT3,
 OP_INT0, OP_INVD, OP_INVLPG, OP_IRET,
 OP_JO, OP_JNO, OP_JC, OP_JNC,
 OP_JZ, OP_JNZ, OP_JBE, OP_JA,
 OP_JS, OP_JNS, OP_JP, OP_JNP,
 OP_JL, OP_JGE, OP_JLE, OP_JG,
 OP_JMP, OP_LAHF, OP_LAR, OP_LDS,
 OP_LES, OP_LFS, OP_LGS, OP_LEA,
 OP_LEAVE, OP_LGDT, OP_LIDT, OP_LLDT,
 OP_PREFIX_LOCK, OP_LODSB, OP_LODSW, OP_LODSB,
 OP_LOOP, OP_LOOPE, OP_LOOPNE, OP_JECXZ,
 OP_LSL, OP_LSS, OP_LTR, OP_MOV,
 OP_MOVSD, OP_MOVSW, OP_MOVSB, OP_MOVSX,
 OP_MOVZX, OP_MUL, OP_NEG, OP_NOP,
 OP_NOT, OP_OR, OP_OUT, OP_OUTSD,
 OP_OUTSW, OP_OUTSB, OP_PUSH, OP_PUSHAD ,
 OP_PUSHFD , OP_POP, OP_POPAD, OP_POPFD ,
 OP_RCL, OP_RCR, OP_RDMSR, OP_RDPMC,
 OP_RDTSC, OP_PREFIX_REPE, OP_PREFIX_REPNB, OP_RETF,
 OP_RETN, OP_ROL, OP_ROR, OP_RSM,
 OP_SAHF, OP_SAR, OP_SBB, OP_SCASB,
 OP_SCASW, OP_SCASB, OP_SETO, OP_SETNO,
 OP_SETC, OP_SETNC, OP_SETZ, OP_SETNZ,
 OP_SETBE, OP_SETA, OP_SETS, OP_SETNS,
 OP_SETP, OP_SETNP, OP_SETL, OP_SETGE,
 OP_SETLE, OP_SETG, OP_SGDT, OP_SIDT,
 OP_SHL, OP_SHLD, OP_SHR, OP_SHRD,
 OP_SLDT, OP_STOSB, OP_STOSW, OP_STOSB,
 OP_STR, OP_STC, OP_STD, OP_STI,
 OP_SUB, OP_SYSCALL, OP_SYSENTER, OP_SYSEXIT,
 OP_SYSRET, OP_TEST, OP_UD2, OP_VERR,
 OP_VERRW, OP_WBINVD, OP_WRMSR, OP_XADD,
 OP_XCHG, OP_XLAT, OP_XOR , OP_FPU,
 OP_F2XM1, OP_FABS, OP_FADD, OP_FADDP,

- OP_FBLD, OP_FBSTP, OP_FCHS, OP_FCLEX,
- OP_FCMOVB, OP_FCMOVBE, OP_FCMOVE, OP_FCMOVNB,
- OP_FCMOVNBE, OP_FCMOVNE, OP_FCMOVNU, OP_FCMOVU,
- OP_FCOM, OP_FCOMI, OP_FCOMIP, OP_FCOMP,
- OP_FCOMPP, OP_FCOS, OP_FDECSTP, OP_FDIV,
- OP_FDIVP, OP_FDIVR, OP_FDIVRP, OP_FFREE,
- OP_FIADD, OP_FICOM, OP_FICOMP, OP_FIDIV,
- OP_FIDIVR, OP_FILD, OP_FIMUL, OP_FINCSTP,
- OP_FINIT, OP_FIST, OP_FISTP, OP_FISTTP,
- OP_FISUB, OP_FISUBR, OP_FLD, OP_FLD1,
- OP_FLDCW, OP_FLDENV, OP_FLDL2E, OP_FLDL2T,
- OP_FLDLG2, OP_FLDLN2, OP_FLDPI, OP_FLDZ,
- OP_FMUL, OP_FMULP, OP_FNOP, OP_FPATAN,
- OP_FPREM, OP_FPREM1, OP_FPTAN, OP_FRNDINT,
- OP_FRSTOR, OP_FSCALE, OP_FSINCOS, OP_FSQRT,
- OP_FSAVE, OP_FST, OP_FSTCW, OP_FSTENV,
- OP_FSTP, OP_FSTSW, OP_FSUB, OP_FSUBP,
- OP_FSUBR, OP_FSUBRP, OP_FTST, OP_FUCOM,
- OP_FUCOMI, OP_FUCOMIP, OP_FUCOMP, OP_FUCOMPP,
- OP_FXAM, OP_FXCH, OP_FXTRACT, OP_FYL2X,
- OP_FYL2XP1 }
- enum DIS_ACCESS {
ACCESS_NOARG, ACCESS_IMM, ACCESS_REL, ACCESS_REG,
ACCESS_MEM }
- enum DIS_SIZE {
SIZEB, SIZEW, SIZED, SIZEF,
SIZEQ, SIZET, SIZEPTR }
- enum X86REGS

7.3.2.1. Detailed Description

7.3.2.2. Enumeration Type Documentation

7.3.2.2.1. enum DIS_ACCESS

Access type

Enumerator:

ACCESS_NOARG arg not present
ACCESS_IMM immediate
ACCESS_REL +/- immediate
ACCESS_REG register
ACCESS_MEM [memory]

7.3.2.2.2. enum **DIS_SIZE** for mem access, immediate and relative

Enumerator:

SIZEB Byte size access
SIZEW Word size access
SIZED Doubleword size access
SIZEF 6-byte access (seg+reg pair)
SIZEQ Quadword access
SIZET 10-byte access
SIZEPTR ptr

7.3.2.2.3. enum **X86OPS**

X86 opcode

Enumerator:

OP_AAA Ascii Adjust after Addition
OP_AAD Ascii Adjust AX before Division
OP_AAM Ascii Adjust AX after Multiply
OP_AAS Ascii Adjust AL after Subtraction
OP_ADD Add
OP_ADC Add with Carry
OP_AND Logical And
OP_ARPL Adjust Requested Privilege Level
OP_BOUND Check Array Index Against Bounds
OP_BSF Bit Scan Forward
OP_BSR Bit Scan Reverse
OP_BSWAP Byte Swap
OP_BT Bit Test
OP BTC Bit Test and Complement
OP BTR Bit Test and Reset
OP BTS Bit Test and Set
OP_CALL Call
OP_CDQ Convert DoubleWord to QuadWord
OP_CWDE Convert Word to DoubleWord
OP_CBW Convert Byte to Word
OP_CLC Clear Carry Flag
OP_CLD Clear Direction Flag
OP CLI Clear Interrupt Flag
OP_CLTS Clear Task-Switched Flag in CR0
OP_CMC Complement Carry Flag
OP_CMOVO Conditional Move if Overflow

OP_CMOVNO Conditional Move if Not Overflow
OP_CMOVC Conditional Move if Carry
OP_CMOVNC Conditional Move if Not Carry
OP_CMOVZ Conditional Move if Zero
OP_CMOVNZ Conditional Move if Non-Zero
OP_CMOVBE Conditional Move if Below or Equal
OP_CMOVA Conditional Move if Above
OP_CMOVS Conditional Move if Sign
OP_CMOVNS Conditional Move if Not Sign
OP_CMOVP Conditional Move if Parity
OP_CMOVNP Conditional Move if Not Parity
OP_CMOVL Conditional Move if Less
OP_CMOVGE Conditional Move if Greater or Equal
OP_CMOVLE Conditional Move if Less than or Equal
OP_CMOVG Conditional Move if Greater
OP_CMP Compare
OP_CMPSD Compare String DoubleWord
OP_CMPSW Compare String Word
OP_CMPSB Compare String Byte
OP_CMPXCHG Compare and Exchange
OP_CMPXCHG8B Compare and Exchange Bytes
OP_CPUID CPU Identification
OP_DAA Decimal Adjust AL after Addition
OP_DAS Decimal Adjust AL after Subtraction
OP_DEC Decrement by 1
OP_DIV Unsigned Divide
OP_ENTER Make Stack Frame for Procedure Parameters
OP_FWAIT Wait
OP_HLT Halt
OP_IDIV Signed Divide
OP_IMUL Signed Multiply
OP_INC Increment by 1
OP_IN INput from port
OP_INSD INput from port to String Doubleword
OP_INSW INput from port to String Word
OP_INSB INput from port to String Byte
OP_INT INTerrupt
OP_INT3 INTerrupt 3 (breakpoint)

OP_INT0 INTerrupt 4 if Overflow
*OP_INV*D Invalidate Internal Caches
OP_INVLPG Invalidate TLB Entry
OP_IRET Interrupt Return
OP_JO Jump if Overflow
OP_JNO Jump if Not Overflow
OP_JC Jump if Carry
OP_JNC Jump if Not Carry
OP_JZ Jump if Zero
OP_JNZ Jump if Not Zero
OP_JBE Jump if Below or Equal
OP_JA Jump if Above
OP_JS Jump if Sign
OP_JNS Jump if Not Sign
OP_JP Jump if Parity
OP_JNP Jump if Not Parity
OP_JL Jump if Less
OP_JGE Jump if Greater or Equal
OP_JLE Jump if Less or Equal
OP_JG Jump if Greater
OP_JMP Jump (unconditional)
OP_LAHF Load Status Flags into AH Register
OP_LAR load Access Rights Byte
OP_LDS Load Far Pointer into DS
OP_LES Load Far Pointer into ES
OP_LFS Load Far Pointer into FS
OP_LGS Load Far Pointer into GS
OP_LEA Load Effective Address
OP_LEAVE High Level Procedure Exit
OP_LGDT Load Global Descript Table Register
OP_LIDT Load Interrupt Descriptor Table Register
OP_LLDT Load Local Descriptor Table Register
OP_PREFIX_LOCK Assert LOCK# Signal Prefix
OP_LODSD Load String Dword
OP_LODSW Load String Word
OP_LODSB Load String Byte
OP_LOOP Loop According to ECX Counter
OP_LOOPE Loop According to ECX Counter and ZF=1

OP_LOOPNE Loop According to ECX Counter and ZF=0
OP_JECXZ Jump if ECX is Zero
OP_LSL Load Segment Limit
OP_LSS Load Far Pointer into SS
OP_LTR Load Task Register
OP_MOV Move
OP_MOVSD Move Data from String to String Doubleword
OP_MOVSW Move Data from String to String Word
OP_MOVSB Move Data from String to String Byte
OP_MOVSX Move with Sign-Extension
OP_MOVZX Move with Zero-Extension
OP_MUL Unsigned Multiply
OP_NEG Two's Complement Negation
OP_NOP No Operation
OP_NOT One's Complement Negation
OP_OR Logical Inclusive OR
OP_OUT Output to Port
OP_OUTSD Output String to Port Doubleword
OP_OUTSW Output String to Port Word
OP_OUTSB Output String to Port Bytes
OP_PUSH Push Onto the Stack
OP_PUSHAD Push All Double General Purpose Registers
OP_PUSHFD Push EFLAGS Register onto the Stack
OP_POP Pop a Value from the Stack
OP_POPAD Pop All Double General Purpose Registers from the Stack
OP_POPFD Pop Stack into EFLAGS Register
OP_RCL Rotate Carry Left
OP_RCR Rotate Carry Right
OP_RDMSR Read from Model Specific Register
OP_RDPMSR Read Performance Monitoring Counters
OP_RDTSC Read Time-Stamp Counter
OP_PREFIX_REPE Repeat String Operation Prefix while Equal
OP_PREFIX_REPNB Repeat String Operation Prefix while Not Equal
OP_RETF Return from Far Procedure
OP_RETN Return from Near Procedure
OP_ROL Rotate Left
OP_ROR Rotate Right

OP_RSM Resume from System Management Mode
OP_SAHF Store AH into Flags
OP_SAR Shift Arithmetic Right
OP_SBB Subtract with Borrow
OP_SCASD Scan String Doubleword
OP_SCASW Scan String Word
OP_SCASB Scan String Byte
OP_SETO Set Byte on Overflow
OP_SETNO Set Byte on Not Overflow
OP_SETC Set Byte on Carry
OP_SETNC Set Byte on Not Carry
OP_SETZ Set Byte on Zero
OP_SETNZ Set Byte on Not Zero
OP_SETBE Set Byte on Below or Equal
OP_SETA Set Byte on Above
OP_SETS Set Byte on Sign
OP_SETNS Set Byte on Not Sign
OP_SETP Set Byte on Parity
OP_SETNP Set Byte on Not Parity
OP_SETL Set Byte on Less
OP_SETGE Set Byte on Greater or Equal
OP_SETLE Set Byte on Less or Equal
OP_SETG Set Byte on Greater
OP_SGDT Store Global Descriptor Table Register
OP_SIDT Store Interrupt Descriptor Table Register
OP_SHL Shift Left
OP_SHLD Double Precision Shift Left
OP_SHR Shift Right
OP_SHRD Double Precision Shift Right
OP_SLDT Store Local Descriptor Table Register
OP_STOSD Store String Doubleword
OP_STOSW Store String Word
OP_STOSB Store String Byte
OP_STR Store Task Register
OP_STC Set Carry Flag
OP_STD Set Direction Flag
OP_STI Set Interrupt Flag
OP_SUB Subtract

OP_SYSCALL Fast System Call
OP_SYSENTER Fast System Call
OP_SYSEXIT Fast Return from Fast System Call
OP_SYSRET Return from Fast System Call
OP_TEST Logical Compare
OP_UD2 Undefined Instruction
OP_VERR Verify a Segment for Reading
OP_VERRW Verify a Segment for Writing
OP_WBINVD Write Back and Invalidate Cache
OP_WRMSR Write to Model Specific Register
OP_XADD Exchange and Add
OP_XCHG Exchange Register/Memory with Register
OP_XLAT Table Look-up Translation
OP_XOR Logical Exclusive OR
OP_FPU FPU operation
OP_F2XM1 Compute $2^x - 1$
OP_FABS Absolute Value
OP_FADD Floating Point Add
OP_FADDP Floating Point Add, Pop
OP_FBLD Load Binary Coded Decimal
OP_FBSTP Store BCD Integer and Pop
OP_FCHS Change Sign
OP_FCLEX Clear Exceptions
OP_FCMOVB Floating Point Move on Below
OP_FCMOVBE Floating Point Move on Below or Equal
OP_FCMOVE Floating Point Move on Equal
OP_FCMOVNB Floating Point Move on Not Below
OP_FCMOVNBE Floating Point Move on Not Below or Equal
OP_FCMOVNE Floating Point Move on Not Equal
OP_FCMOVNU Floating Point Move on Not Unordered
OP_FCMOVU Floating Point Move on Unordered
OP_FCOM Compare Floating Pointer Values and Set FPU Flags
OP_FCOMI Compare Floating Pointer Values and Set EFLAGS
OP_FCOMIP Compare Floating Pointer Values and Set EFLAGS, Pop
OP_FCOMP Compare Floating Pointer Values and Set FPU Flags,
 Pop
OP_FCOMPP Compare Floating Pointer Values and Set FPU Flags,
 Pop Twice
OP_FCOS Cosine

OP_FDECSTP Decrement Stack Top Pointer
OP_FDIV Floating Point Divide
OP_FDIVP Floating Point Divide, Pop
OP_FDIVR Floating Point Reverse Divide
OP_FDIVRP Floating Point Reverse Divide, Pop
OP_FFREE Free Floating Point Register
OP_FIADD Floating Point Add
OP_FICOM Compare Integer
OP_FICOMP Compare Integer, Pop
OP_FIDIV Floating Point Divide by Integer
OP_FIDIVR Floating Point Reverse Divide by Integer
OP_FILD Load Integer
OP_FIMUL Floating Point Multiply with Integer
OP_FINCSTP Increment Stack-Top Pointer
OP_FINIT Initialize Floating-Point Unit
OP_FIST Store Integer
OP_FISTP Store Integer, Pop
OP_FISTTP Store Integer with Truncation
OP_FISUB Floating Point Integer Subtract
OP_FISUBR Floating Point Reverse Integer Subtract
OP_FLD Load Floating Point Value
OP_FLD1 Load Constant 1
OP_FLDCW Load x87 FPU Control Word
OP_FLDENV Load x87 FPU Environment
OP_FLDL2E Load Constant $\log_2(e)$
OP_FLDL2T Load Constant $\log_2(10)$
OP_FLDLG2 Load Constant $\log_{10}(2)$
OP_FLDLN2 Load Constant $\log_e(2)$
OP_FLDPI Load Constant PI
OP_FLDZ Load Constant Zero
OP_FMUL Floating Point Multiply
OP_FMULP Floating Point Multiply, Pop
OP_FNOP No Operation
OP_FPATAN Partial Arctangent
OP_FPREM Partial Remainder
OP_FPREM1 Partial Remainder
OP_FPTAN Partial Tangent
OP_FRNDINT Round to Integer

OP_FRSTOR Restore x86 FPU State
OP_FSCALE Scale
OP_FSINCOS Sine and Cosine
OP_FSQRT Square Root
OP_FSAVE Store x87 FPU State
OP_FST Store Floating Point Value
OP_FSTCW Store x87 FPU Control Word
OP_FSTENV Store x87 FPU Environment
OP_FSTP Store Floating Point Value, Pop
OP_FSTSW Store x87 FPU Status Word
OP_FSUB Floating Point Subtract
OP_FSUBP Floating Point Subtract, Pop
OP_FSUBR Floating Point Reverse Subtract
OP_FSUBRP Floating Point Reverse Subtract, Pop
OP_FTST Floating Point Test
OP_FUCOM Floating Point Unordered Compare
OP_FUCOMI Floating Point Unordered Compare with Integer
OP_FUCOMIP Floating Point Unorder Compare with Integer, Pop
OP_FUCOMP Floating Point Unorder Compare, Pop
OP_FUCOMPP Floating Point Unorder Compare, Pop Twice
OP_FXAM Examine ModR/M
OP_FXCH Exchange Register Contents
OP_FXTRACT Extract Exponent and Significand
OP_FYL2X Compute $y \cdot \log_2 x$
OP_FYL2XP1 Compute $y \cdot \log_2(x+1)$

7.3.2.2.4. enum X86REGS

X86 registers

7.3.3. bytecode__execs.h File Reference

Data Structures

- struct [cli_exe_section](#)
- struct [cli_exe_info](#)

7.3.3.1. Detailed Description

7.3.4. bytecode__pe.h File Reference

Data Structures

- struct [pe_image_file_hdr](#)
- struct [pe_image_data_dir](#)
- struct [pe_image_optional_hdr32](#)
- struct [pe_image_optional_hdr64](#)
- struct [pe_image_section_hdr](#)
- struct [cli_pe_hook_data](#)

7.3.4.1. Detailed Description

7.4. High level API

7.4.1. bytecode_local.h File Reference

Data Structures

- struct [DIS_mem_arg](#)
- struct [DIS_arg](#)
- struct [DIS_fixed](#)

Defines

- `#define VIRUSNAME_PREFIX(name) const char __clambc_virusname_prefix[] = name;`
- `#define VIRUSNAMES(...) const char *const __clambc_virusnames[] = {__VA_ARGS__};`
- `#define PE_UNPACKER_DECLARE const uint16_t __clambc_kind = BC_PE_UNPACKER;`
- `#define PDF_HOOK_DECLARE const uint16_t __clambc_kind = BC_PDF;`
- `#define SIGNATURES_DECL_BEGIN struct __Signatures {`
- `#define DECLARE_SIGNATURE(name)`
- `#define SIGNATURES_DECL_END };`
- `#define TARGET(tgt) const unsigned short __Target = (tgt);`
- `#define COPYRIGHT(c) const char *const __Copyright = (c);`
- `#define ICONGROUP1(group) const char *const __IconGroup1 = (group);`
- `#define ICONGROUP2(group) const char *const __IconGroup2 = (group);`
- `#define FUNCTIONALITY_LEVEL_MIN(m) const unsigned short __FuncMin = (m);`
- `#define FUNCTIONALITY_LEVEL_MAX(m) const unsigned short __FuncMax = (m);`
- `#define SIGNATURES_DEF_BEGIN`
- `#define DEFINE_SIGNATURE(name, hex)`
- `#define SIGNATURES_END };\`

Functions

- static force_inline void overloadable_func [debug](#) (const char *str)
- static force_inline void overloadable_func [debug](#) (const uint8_t *str)
- static force_inline void overloadable_func [debug](#) (uint32_t a)
- void [debug](#) (...) [__attribute__\(\(overloadable\)\)](#)
- static force_inline uint32_t [count_match](#) (__Signature sig)
- static force_inline uint32_t [matches](#) (__Signature sig)
- static force_inline uint32_t [match_location](#) (__Signature sig, uint32_t goback)

- static force_inline uint32_t `match_location_check` (___Signature sig, uint32_t goback, const char *static_start, uint32_t static_len)
- static force_inline overloadable_func void `foundVirus` (const char *virusname)
- static force_inline void overloadable_func `foundVirus` (void)
- static force_inline uint32_t `getFileSize` (void)
- bool `__is_bigendian` (void) `__attribute__((const)) __attribute__((nothrow))`
- static uint32_t force_inline `le32_to_host` (uint32_t v)
- static uint64_t force_inline `le64_to_host` (uint64_t v)
- static uint16_t force_inline `le16_to_host` (uint16_t v)
- static uint32_t force_inline `cli_readint32` (const void *buff)
- static uint16_t force_inline `cli_readint16` (const void *buff)
- static void force_inline `cli_writeint32` (void *offset, uint32_t v)
- static force_inline bool `hasExeInfo` (void)
- static force_inline bool `isPE64` (void)
- static static force_inline force_inline uint8_t `getPEMajorLinkerVersion` (void)
- static force_inline uint8_t `getPEMinorLinkerVersion` (void)
- static force_inline uint32_t `getPESizeOfCode` (void)
- static force_inline uint32_t `getPESizeOfInitializedData` (void)
- static force_inline uint32_t `getPESizeOfUninitializedData` (void)
- static force_inline uint32_t `getPEBaseOfCode` (void)
- static force_inline uint32_t `getPEBaseOfData` (void)
- static force_inline uint64_t `getPEImageBase` (void)
- static force_inline uint32_t `getPESectionAlignment` (void)
- static force_inline uint32_t `getPEFileAlignment` (void)
- static force_inline uint16_t `getPEMajorOperatingSystemVersion` (void)
- static force_inline uint16_t `getPEMinorOperatingSystemVersion` (void)
- static force_inline uint16_t `getPEMajorImageVersion` (void)
- static force_inline uint16_t `getPEMinorImageVersion` (void)
- static force_inline uint16_t `getPEMajorSubsystemVersion` (void)
- static force_inline uint16_t `getPEMinorSubsystemVersion` (void)
- static force_inline uint32_t `getPEWin32VersionValue` (void)
- static force_inline uint32_t `getPESizeOfImage` (void)
- static force_inline uint32_t `getPESizeOfHeaders` (void)
- static force_inline uint32_t `getPEChecksum` (void)
- static force_inline uint16_t `getPESubsystem` (void)
- static force_inline uint16_t `getPEDllCharacteristics` (void)

Return the PE DllCharacteristics.

- static force_inline uint32_t `getPESizeOfStackReserve` (void)
- static force_inline uint32_t `getPESizeOfStackCommit` (void)
- static force_inline uint32_t `getPESizeOfHeapReserve` (void)
- static force_inline uint32_t `getPESizeOfHeapCommit` (void)
- static force_inline uint32_t `getPELoaderFlags` (void)
- static force_inline uint16_t `getPEMachine` ()

- static force_inline uint32_t [getPETimeDateStamp](#) ()
- static force_inline uint32_t [getPEPointerToSymbolTable](#) ()
- static force_inline uint32_t [getPENumberOfSymbols](#) ()
- static force_inline uint16_t [getPESizeOfOptionalHeader](#) ()
- static force_inline uint16_t [getPECharacteristics](#) ()
- static force_inline bool [getPEisDLL](#) ()
- static force_inline uint32_t [getPEDataDirRVA](#) (unsigned n)
- static force_inline uint32_t [getPEDataDirSize](#) (unsigned n)
- static force_inline uint16_t [getNumberOfSections](#) (void)
- static uint32_t [getPELFANew](#) (void)
- static force_inline int [readPESectionName](#) (unsigned char name[8], unsigned n)
- static force_inline uint32_t [getEntryPoint](#) (void)
- static force_inline uint32_t [getExeOffset](#) (void)
- static force_inline uint32_t [getImageBase](#) (void)
- static uint32_t [getVirtualEntryPoint](#) (void)
- static uint32_t [getSectionRVA](#) (unsigned i)
- static uint32_t [getSectionVirtualSize](#) (unsigned i)
- static force_inline bool [readRVA](#) (uint32_t rva, void *buf, size_t bufsize)
- static void * [memchr](#) (const void *s, int c, size_t n)
- void * [memset](#) (void *src, int c, uintptr_t n) [__attribute__\(\(nothrow\)\)](#) [__attribute__\(\(nonnull\(\(1\)\)\)\)](#)
- void * [memmove](#) (void *dst, const void *src, uintptr_t n) [__attribute__\(\(nothrow\)\)](#) [__attribute__\(\(nonnull\(1\)\)\)](#)
- void *void * [memcpy](#) (void *restrict dst, const void *restrict src, uintptr_t n) [__attribute__\(\(nothrow\)\)](#) [__attribute__\(\(nonnull\(1\)\)\)](#)
- void *void *int [memcmp](#) (const void *s1, const void *s2, uint32_t n) [__attribute__\(\(nothrow\)\)](#) [__attribute__\(\(pure\)\)](#) [__attribute__\(\(nonnull\(1\)\)\)](#)
- static force_inline uint32_t [DisassembleAt](#) (struct [DIS_fixed](#) *result, uint32_t offset, uint32_t len)
- static int32_t [ilog2_compat](#) (uint32_t a, uint32_t b)

Variables

- void *void *int struct [DIS_mem_arg](#) [__attribute__\(\(nonnull\)\)](#)

7.4.1.1. Detailed Description

7.4.1.2. Define Documentation

7.4.1.2.1. #define COPYRIGHT(*c*) const char *const ____Copyright = (c); Defines an alternative copyright for this bytecode.

[config](#)

This will also prevent the sourcecode from being embedded into the bytecode

7.4.1.2.2. #define DECLARE_SIGNATURE(*name*) Value:

```
const char *name##_sig;\n    __Signature name;
```

Declares a name for a subsignature.

[config](#)

7.4.1.2.3. #define DEFINE_SIGNATURE(*name*, *hex*) Value:

```
.name##_sig = (hex),\n    .name = {__COUNTER__ - __signature_bias},
```

Defines the pattern for a previously declared subsignature.

See also

[DECLARE_SIGNATURE](#)

[config](#)

Parameters

name the name of a previously declared subsignature

hex the pattern for this subsignature

7.4.1.2.4. #define FUNCTIONALITY_LEVEL_MAX(*m*) const unsigned short __FuncMax = (m); Define the

maximum engine functionality level required for this bytecode/logical signature. Engines newer than this will skip loading the bytecode. You can use the 'enum FunctionalityLevels' constants here.

[config](#)

7.4.1.2.5. #define FUNCTIONALITY_LEVEL_MIN(*m*) const unsigned short __FuncMin = (m); Define the

minimum engine functionality level required for this bytecode/logical signature. Engines older than this will skip loading the bytecode. You can use the 'enum FunctionalityLevels' constants here.

[config](#)

7.4.1.2.6. #define ICONGROUP1(*group*) const char *const __-IconGroup1 = (group); Define IconGroup1 for logical

signature.

See logical signature documentation for what it is

[config](#)

7.4.1.2.7. #define ICONGROUP2(*group*) const char *const ____-IconGroup2 = (group); Define IconGroup2 for logical signature. See logical signature documentation for what it is.

[config](#)

7.4.1.2.8. #define PDF_HOOK_DECLARE const uint16_t ____-clambc_kind = BC_PDF; Make the current bytecode a PDF hook. Having a logical signature doesn't make sense here, since logical signature is evaluated AFTER these hooks run.

[config](#)

This hook is called several times, use [pdf_get_phase\(\)](#) to find out in which phase you got called.

7.4.1.2.9. #define PE_UNPACKER_DECLARE const uint16_t ____-clambc_kind = BC_PE_UNPACKER; Make the current bytecode a PE hook, i.e. it will be called once the logical signature trigger matches (or always if there is none), and you have access to all the PE information. By default you only have access to `execs.h` information, and not to PE field information (even for PE files).

[config](#)

7.4.1.2.10. #define SIGNATURES_DECL_BEGIN struct ____-Signatures { Marks the beginning of the subsignature name declaration section.

[config](#)

7.4.1.2.11. #define SIGNATURES_DECL_END }; Marks the end of the subsignature name declaration section.

[config](#)

7.4.1.2.12. #define SIGNATURES_DEF_BEGIN **Value:**

```
static const unsigned __signature_bias = __COUNTER__ + 1;\nconst struct __Signatures Signatures = {\n
```

Marks the beginning of subsignature pattern definitions.

[config](#)

See also

[SIGNATURES_DECL_BEGIN](#)

7.4.1.2.13. #define SIGNATURES__END ; Marks the end of the subsignature pattern definitions.

[config](#)

7.4.1.2.14. #define TARGET(*tgt*) const unsigned short __Target = (*tgt*); Defines the ClamAV file target.

[config](#)

Parameters

tgt ClamAV signature type (0 - raw, 1 - PE, etc.)

7.4.1.2.15. #define VIRUSNAME__PREFIX(*name*) const char __clambc_virusname_prefix[] = *name*; Declares the virusname prefix.

[config](#)

Parameters

name the prefix common to all viruses reported by this bytecode

7.4.1.2.16. #define VIRUSNAMES(...) const char *const __clambc_virusnames[] = {__VA_ARGS__}; Declares all the virusnames that this bytecode can report.

[config](#)

Parameters

... a comma-separated list of strings interpreted as virusnames

7.4.1.3. Function Documentation

7.4.1.3.1. bool __is_bigendian (void) const Returns true if the bytecode is executing on a big-endian CPU.

Returns

true if executing on bigendian CPU, false otherwise

Environment

This will be optimized away in libclamav, but it must be used when dealing with endianness for portability reasons. For example whenever you read a 32-bit integer from a file, it can be written in little-endian convention (x86 CPU for

example), or big-endian convention (PowerPC CPU for example). If the file always contains little-endian integers, then conversion might be needed. ClamAV bytecodes by their nature must only handle known-endian integers, if endianness can change, then both situations must be taken into account (based on a 1-byte field for example).

7.4.1.3.2. static uint16_t force_inline cli_readint16 (const void * *buff*) [static] Reads from the specified buffer a 16-bit of little-endian integer.

Data structure

Parameters

[in] *buff* pointer to buffer

Returns

16-bit little-endian integer converted to host endianness

7.4.1.3.3. static uint32_t force_inline cli_readint32 (const void * *buff*) [static] Reads from the specified buffer a 32-bit of little-endian integer.

Data structure

Parameters

[in] *buff* pointer to buffer

Returns

32-bit little-endian integer converted to host endianness

7.4.1.3.4. static void force_inline cli_writeint32 (void * *offset*, uint32_t *v*) [static] Writes the specified value into the specified buffer in little-endian order

Data structure

Parameters

[out] *offset* pointer to buffer to write to

[in] *v* value to write

7.4.1.3.5. static force_inline uint32_t count_match (__Signature *sig*) [static] Returns how many times the specified signature matched.

Parameters

sig name of subsignature queried

Returns

number of times this subsignature matched in the entire file

Engine query

This is a constant-time operation, the counts for all subsignatures are already computed.

7.4.1.3.6. void debug (...) debug is an overloaded function (yes clang supports that in C!), but it only works on strings, and integers. Give an error on any other type

7.4.1.3.7. static force_inline void overloadable_func debug (const char * *str*) [static] Prints *str* to clamscan's --debug output.

Parameters

str null terminated string

7.4.1.3.8. static force_inline void overloadable_func debug (const uint8_t * *str*) [static] Prints *str* to clamscan's --debug output.

Parameters

str null terminated string

7.4.1.3.9. static force_inline void overloadable_func debug (uint32_t *a*) [static] Prints *a* integer to clamscan's --debug output.

Parameters

a integer

7.4.1.3.10. static force_inline uint32_t DisassembleAt (struct DIS__fixed * *result*, uint32_t *offset*, uint32_t *len*) [static] Disassembles one X86 instruction starting at the specified offset.

Disassemble**Parameters**

[out] *result* disassembly result

[in] *offset* start disassembling from this offset, in the current file

[in] *len* max amount of bytes to disassemble

Returns

offset where disassembly ended

7.4.1.3.11. static force_inline overloadable_func void foundVirus (const char * *virusname*) [static] Sets the specified virusname as the virus detected by this bytecode.

Scan

Parameters

virusname the name of the virus, excluding the prefix, must be one of the virusnames declared in VIRUSNAMES.

See also

[VIRUSNAMES](#)

7.4.1.3.12. static force_inline void overloadable_func foundVirus (void) [static] Like [foundVirus\(\)](#) but just use the prefix as virusname

7.4.1.3.13. static force_inline uint32_t getEntryPoint (void) [static] Returns the offset of the EntryPoint in the executable file.

PE

Returns

offset of EP as 32-bit unsigned integer

7.4.1.3.14. static force_inline uint32_t getExeOffset (void) [static] Returns the offset of the executable in the file.

PE

Returns

offset of embedded executable inside file.

7.4.1.3.15. static force_inline uint32_t getFilesize (void) [static] Returns the currently scanned file's size.

File operation

Returns

file size as 32-bit unsigned integer

7.4.1.3.16. static force_inline uint32_t getImageBase (void) [static] Returns the ImageBase with the correct endian conversion. Only works if the bytecode is a PE hook (i.e. you invoked PE_UNPACKER_DECLARE)

PE**Returns**

ImageBase of PE file, 0 - for non-PE hook

7.4.1.3.17. static force__inline uint16__t getNumberOfSections (void) [static] Returns the number of sections in this executable file.

PE**Returns**

number of sections as 16-bit unsigned integer

7.4.1.3.18. static force__inline uint32__t getPEBaseOfCode (void) [static] Return the PE BaseOfCode.

PE**Returns**

PE BaseOfCode, or 0 if not in PE hook.

7.4.1.3.19. static force__inline uint32__t getPEBaseOfData (void) [static] Return the PE BaseOfData.

PE**Returns**

PE BaseOfData, or 0 if not in PE hook.

7.4.1.3.20. static force__inline uint16__t getPECharacteristics () [static] Returns PE characteristics. For example you can use this to check whether it is a DLL (0x2000).

PE**Returns**

characteristic of PE file, or 0 if not in PE hook

7.4.1.3.21. static force__inline uint32__t getPEChecksum (void) [static] Return the PE CheckSum.

PE**Returns**

PE CheckSum, or 0 if not in PE hook

7.4.1.3.22. static force__inline uint32_t getPEDirRVA (unsigned n) [static] Gets the virtual address of specified image data directory.

PE

Parameters

n image directory requested

Returns

Virtual Address of requested image directory

7.4.1.3.23. static force__inline uint32_t getPEDirSize (unsigned n) [static] Gets the size of the specified image data directory.

PE

Parameters

n image directory requested

Returns

Size of requested image directory

7.4.1.3.24. static force__inline uint16_t getPEDllCharacteristics (void) [static]
Return the PE DllCharacteristics.

PE

Returns

PE DllCharacteristics, or 0 if not in PE hook

7.4.1.3.25. static force__inline uint32_t getPEFileAlignment (void) [static] Return the PE FileAlignment.

PE

Returns

PE FileAlignment, or 0 if not in PE hook

7.4.1.3.26. static force__inline uint64_t getPEImageBase (void) [static] Return the PE ImageBase as 64-bit integer.

PE

Returns

PE ImageBase as 64-bit int, or 0 if not in PE hook

7.4.1.3.27. static force__inline bool getPEisDLL () [static] Returns whether this is a DLL. Use this only in a PE hook!

PE

Returns

true - the file is a DLL false - file is not a DLL

7.4.1.3.28. static uint32_t getPELFANew (void) [static] Gets the offset to the PE header.

PE

Returns

offset to the PE header, or 0 if not in PE hook

7.4.1.3.29. static force__inline uint32_t getPELoaderFlags (void) [static] Return the PE LoaderFlags.

PE

Returns

PE LoaderFlags or 0 if not in PE hook

7.4.1.3.30. static force__inline uint16_t getPEMachine () [static]
Returns the CPU this executable runs on, see libclamav/pe.c for possible values.

PE

Returns

PE Machine or 0 if not in PE hook

7.4.1.3.31. static force__inline uint16_t getPEMajorImageVersion (void) [static] Return the PE MajorImageVersion.

PE

Returns

PE MajorImageVersion, or 0 if not in PE hook

7.4.1.3.32. static force__inline force__inline uint8__t getPEMajorLinkerVersion (void) [static] Returns MajorLinkerVersion for this PE file.

PE

Returns

PE MajorLinkerVersion or 0 if not in PE hook

7.4.1.3.33. static force__inline uint16__t getPEMajorOperatingSystemVersion (void) [static] Return the PE MajorOperatingSystemVersion.

PE

Returns

PE MajorOperatingSystemVersion, or 0 if not in PE hook

7.4.1.3.34. static force__inline uint16__t getPEMajorSubsystemVersion (void) [static] Return the PE MajorSubsystemVersion.

PE

Returns

PE MajorSubsystemVersion or 0 if not in PE hook

7.4.1.3.35. static force__inline uint16__t getPEMinorImageVersion (void) [static] Return the PE MinorImageVersion.

PE

Returns

PE MinorImageVersion, or 0 if not in PE hook

7.4.1.3.36. static force__inline uint8__t getPEMinorLinkerVersion (void) [static] Returns MinorLinkerVersion for this PE file.

PE

Returns

PE MinorLinkerVersion or 0 if not in PE hook

7.4.1.3.37. static force_inline uint16_t getPEMinorOperatingSystemVersion (void) [static] Return the PE MinorOperatingSystemVersion.

PE

Returns

PE MinorOperatingSystemVersion, or 0 if not in PE hook

7.4.1.3.38. static force_inline uint16_t getPEMinorSubsystemVersion (void) [static] Return the PE MinorSubsystemVersion.

PE

Returns

PE MinorSubsystemVersion, or 0 if not in PE hook

7.4.1.3.39. static force_inline uint32_t getPENumberOfSymbols () [static] Returns the PE number of debug symbols

PE

Returns

PE NumberOfSymbols or 0 if not in PE hook

7.4.1.3.40. static force_inline uint32_t getPEPointerToSymbolTable () [static] Returns pointer to the PE debug symbol table

PE

Returns

PE PointerToSymbolTable or 0 if not in PE hook

7.4.1.3.41. static force_inline uint32_t getPESectionAlignment (void) [static] Return the PE SectionAlignment.

PE

Returns

PE SectionAlignment, or 0 if not in PE hook

7.4.1.3.42. static force__inline uint32__t getPESizeOfCode (void)
[static] Return the PE SizeOfCode.

PE

Returns

PE SizeOfCode or 0 if not in PE hook

7.4.1.3.43. static force__inline uint32__t getPESizeOfHeaders (void)
[static] Return the PE SizeOfHeaders.

PE

Returns

PE SizeOfHeaders, or 0 if not in PE hook

7.4.1.3.44. static force__inline uint32__t getPESizeOfHeapCommit (void)
[static] Return the PE SizeOfHeapCommit.

PE

Returns

PE SizeOfHeapCommit, or 0 if not in PE hook

7.4.1.3.45. static force__inline uint32__t getPESizeOfHeapReserve (void)
[static] Return the PE SizeOfHeapReserve.

PE

Returns

PE SizeOfHeapReserve, or 0 if not in PE hook

7.4.1.3.46. static force__inline uint32__t getPESizeOfImage (void)
[static] Return the PE SizeOfImage.

PE

Returns

PE SizeOfImage, or 0 if not in PE hook

7.4.1.3.47. static force__inline uint32__t getPESizeOfInitializedData (void)
[static] Return the PE SizeofInitializedData.

PE

Returns

PE SizeOfInitializeData or 0 if not in PE hook

7.4.1.3.48. static force__inline uint16__t getPESizeOfOptionalHeader () [static] Returns the size of PE optional header.

PE

Returns

size of PE optional header, or 0 if not in PE hook

7.4.1.3.49. static force__inline uint32__t getPESizeOfStackCommit (void) [static] Return the PE SizeOfStackCommit.

PE

Returns

PE SizeOfStackCommit, or 0 if not in PE hook

7.4.1.3.50. static force__inline uint32__t getPESizeOfStackReserve (void) [static] Return the PE SizeOfStackReserve.

PE

Returns

PE SizeOfStackReserve, or 0 if not in PE hook

7.4.1.3.51. static force__inline uint32__t getPESizeOfUninitialized-Data (void) [static] Return the PE SizeofUninitializedData.

PE

Returns

PE SizeofUninitializedData or 0 if not in PE hook

7.4.1.3.52. static force__inline uint16__t getPESubsystem (void) [static] Return the PE Subsystem.

PE

Returns

PE subsystem, or 0 if not in PE hook

7.4.1.3.53. static force__inline uint32_t getPETimeStamp ()
[static] Returns the PE TimeDateStamp from headers

PE

Returns

PE TimeDateStamp or 0 if not in PE hook

7.4.1.3.54. static force__inline uint32_t getPEWin32VersionValue (
void) [static] Return the PE Win32VersionValue.

PE

Returns

PE Win32VersionValue, or 0 if not in PE hook

7.4.1.3.55. static uint32_t getSectionRVA (unsigned i) [static]
 Return the RVA of the specified section

PE

Parameters

i section index (from 0)

Returns

RVA of section, or -1 if invalid

7.4.1.3.56. static uint32_t getSectionVirtualSize (unsigned i)
[static] Return the virtual size of the specified section.

PE

Parameters

i section index (from 0)

Returns

VSZ of section, or -1 if invalid

7.4.1.3.57. static uint32_t getVirtualEntryPoint (void) [static]
 The address of the EntryPoint. Use this for matching EP against sections.

PE

Returns

virtual address of EntryPoint, or 0 if not in PE hook

7.4.1.3.58. static force_inline bool hasExeInfo (void) [static]

Returns whether the current file has executable information.

[PE](#)

Returns

true if the file has exe info, false otherwise

7.4.1.3.59. static int32_t ilog2_compat (uint32_t a, uint32_t b) [inline, static] `ilog2_compat` for 0.96 compatibility, you should use [ilog2\(\)](#) 0.96.1 API instead of this one!

7.4.1.3.60. static force_inline bool isPE64 (void) [static] Returns whether this is a PE32+ executable.

[PE](#)

Returns

true if this is a PE32+ executable

7.4.1.3.61. static uint16_t force_inline le16_to_host (uint16_t v) [static] Converts the specified value if needed, knowing it is in little endian order.

[Data structure](#)

Parameters

[in] *v* 16-bit integer as read from a file

Returns

integer converted to host's endianness

7.4.1.3.62. static uint32_t force_inline le32_to_host (uint32_t v) [static] Converts the specified value if needed, knowing it is in little endian order.

[Data structure](#)

Parameters

[in] *v* 32-bit integer as read from a file

Returns

integer converted to host's endianness

7.4.1.3.63. static uint64_t force_inline le64_to_host (uint64_t v) [static] Converts the specified value if needed, knowing it is in little endian order.

Data structure

Parameters

[in] *v* 64-bit integer as read from a file

Returns

integer converted to host's endianness

7.4.1.3.64. static force_inline uint32_t match_location (____-Signature *sig*, uint32_t *goback*) [static] Returns the offset of the match.

Engine query

Parameters

sig - Signature

goback - max length of signature

Returns

offset of match

7.4.1.3.65. static force_inline int32_t match_location_check (____-Signature *sig*, uint32_t *goback*, const char * *static_start*, uint32_t *static_len*) [static] Like [match_location\(\)](#), but also checks that the match starts with the specified hex string.

Engine query

It is recommended to use this for safety and compatibility with 0.96.1

Parameters

sig - signature

goback - maximum length of signature (till start of last subsig)

static_start - static string that sig must begin with

static_len - static string that sig must begin with - length

Returns

≥ 0 - offset of match -1 - no match

7.4.1.3.66. static force_inline uint32_t matches (____Signature *sig*) [**static**] Returns whether the specified subsignature has matched at least once.

Engine query

Parameters

sig name of subsignature queried

Returns

1 if subsignature one or more times, 0 otherwise

7.4.1.3.67. static void* memchr (const void * *s*, int *c*, size_t *n*) [**static**] Scan the first *n* bytes of the buffer *s*, for the character *c*.

String operation

Parameters

[in] *s* buffer to scan

c character to look for

n size of buffer

Returns

a pointer to the first byte to match, or NULL if not found.

7.4.1.3.68. void* void* int memcmp (const void * *s1*, const void * *s2*, uint32_t *n*) Compares two memory buffers.

String operation

Parameters

[in] *s1* buffer one

[in] *s2* buffer two

[in] *n* amount of bytes to copy

Returns

an integer less than, equal to, or greater than zero if the first *n* bytes of *s1* are found, respectively, to be less than, to match, or be greater than the first *n* bytes of *s2*.

7.4.1.3.69. void* void* memcpy (void *restrict *dst*, const void *restrict *src*, uintptr_t *n*) Copies data between two non-overlapping buffers.

String operation

Parameters

- [out] *dst* destination buffer
- [in] *src* source buffer
- [in] *n* amount of bytes to copy

Returns

dst

7.4.1.3.70. void* memmove (void * *dst*, const void * *src*, uintptr_t *n*) Copies data between two possibly overlapping buffers.

String operation

Parameters

- [out] *dst* destination buffer
- [in] *src* source buffer
- [in] *n* amount of bytes to copy

Returns

dst

7.4.1.3.71. void* memset (void * *src*, int *c*, uintptr_t *n*) Fills the specified buffer to the specified value.

String operation

Parameters

- [out] *src* pointer to buffer
- [in] *c* character to fill buffer with
- [in] *n* length of buffer

Returns

src

7.4.1.3.72. `static force_inline int readPESectionName (unsigned char name[8], unsigned n) [static]` Read name of requested PE section.

[PE](#)

Parameters

[out] *name* name of PE section

[in] *n* PE section requested

Returns

0 if successful, <0 otherwise

7.4.1.3.73. `static force_inline bool readRVA (uint32_t rva, void * buf, size_t bufsize) [static]` read the specified amount of bytes from the PE file, starting at the address specified by RVA.

[PE](#)

Parameters

rva the Relative Virtual Address you want to read from (will be converted to file offset)

[out] *buf* destination buffer

bufsize size of buffer

Returns

true on success (full read), false on any failure

7.4.1.4. Variable Documentation

7.4.1.4.1. `union unaligned_16 __attribute__` disassembled memory operand: `scale_reg*scale + add_reg + displacement`

[Disassemble](#)

CHAPTER 8

Copyright and License

8.1. The ClamAV Bytecode Compiler

The ClamAV Bytecode Compiler is released under the GNU General Public License version 2.

The following directories are under the GNU General Public License version 2: ClamBC, docs, driver, editor, examples, ifacegen.

Copyright (C) 2009 Sourcefire, Inc.

This program is free software; you can redistribute it and/or modify it under the terms of the GNU General Public License version 2 as published by the Free Software Foundation.

This program is distributed in the hope that it will be useful, but WITHOUT ANY WARRANTY; without even the implied warranty of MERCHANTABILITY or FITNESS FOR A PARTICULAR PURPOSE. See the GNU General Public License for more details.

You should have received a copy of the GNU General Public License along with this program; if not, write to the Free Software Foundation, Inc., 51 Franklin Street, Fifth Floor, Boston, MA 02110-1301, USA.

It uses the LLVM compiler framework, contained in the following directories: llvm, clang. They have this copyright:

```
=====
LLVM Release License
=====
```

```
University of Illinois/NCSA
Open Source License
```

```
Copyright (c) 2003-2009 University of Illinois at Urbana-Champaign.
All rights reserved.
```

```
Developed by:
```

```
    LLVM Team
```

```
    University of Illinois at Urbana-Champaign
```

```
    http://llvm.org
```

```
Permission is hereby granted, free of charge, to any person obtaining a copy of
```


this software and associated documentation files (the "Software"), to deal with the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

- * Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimers.
- * Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimers in the documentation and/or other materials provided with the distribution.
- * Neither the names of the LLVM Team, University of Illinois at Urbana-Champaign, nor the names of its contributors may be used to endorse or promote products derived from this Software without specific prior written permission.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE CONTRIBUTORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS WITH THE SOFTWARE.

=====
Copyrights and Licenses for Third Party Software Distributed with LLVM:
=====

The LLVM software contains code written by third parties. Such software will have its own individual LICENSE.TXT file in the directory in which it appears. This file will describe the copyrights, license, and restrictions which apply to that code.

The disclaimer of warranty in the University of Illinois Open Source License applies to all code in the LLVM Distribution, and nothing in any of the other licenses gives permission to use the names of the LLVM Team or the University of Illinois to endorse or promote products derived from this Software.

The following pieces of software have additional or alternate copyrights, licenses, and/or restrictions:

Program	Directory
-----	-----
Autoconf	llvm/autoconf
	llvm/projects/ModuleMaker/autoconf
	llvm/projects/sample/autoconf
CellSPU backend	llvm/lib/Target/CellSPU/README.txt
Google Test	llvm/utils/unittest/googletest
OpenBSD regex	llvm/lib/Support/{reg*, COPYRIGHT.regex}

It also uses re2c, contained in driver/clamdriver/re2c. This code is public domain:

Originally written by Peter Bumbulis (peter@csg.uwaterloo.ca)

Currently maintained by:

- * Dan Nuffer <nuffer@users.sourceforge.net>
- * Marcus Boerger <helly@users.sourceforge.net>
- * Hartmut Kaiser <hkaiser@users.sourceforge.net>

The re2c distribution can be found at:

<http://sourceforge.net/projects/re2c/>

re2c is distributed with no warranty whatever. The code is certain to contain errors. Neither the author nor any contributor takes responsibility for any consequences of its use.

re2c is in the public domain. The data structures and algorithms used in re2c are all either taken from documents available to the general public or are inventions of the author. Programs generated by re2c may be distributed freely. re2c itself may be distributed freely, in source or binary, unchanged or modified. Distributors may charge whatever fees they can obtain for re2c.

If you do make use of re2c, or incorporate it into a larger project an acknowledgement somewhere (documentation, research report, etc.) would be appreciated.

8.2. Bytecode

The headers used when compiling bytecode have these license (`clang/lib/Headers/{bcfeatures,bytecode*}.h`):

Copyright (C) 2009 Sourcefire, Inc.
All rights reserved.

Redistribution and use in source and binary forms, with or without modification, are permitted provided that the following conditions are met:

1. Redistributions of source code must retain the above copyright notice, this list of conditions and the following disclaimer.
2. Redistributions in binary form must reproduce the above copyright notice, this list of conditions and the following disclaimer in the documentation and/or other materials provided with the distribution.

THIS SOFTWARE IS PROVIDED BY THE REGENTS AND CONTRIBUTORS ‘‘AS IS’’ AND ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE ARE DISCLAIMED. IN NO EVENT SHALL THE REGENTS OR CONTRIBUTORS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL, EXEMPLARY, OR CONSEQUENTIAL DAMAGES (INCLUDING, BUT NOT LIMITED TO, PROCUREMENT OF SUBSTITUTE GOODS OR SERVICES; LOSS OF USE, DATA, OR PROFITS; OR BUSINESS INTERRUPTION) HOWEVER CAUSED AND ON ANY THEORY OF LIABILITY, WHETHER IN CONTRACT, STRICT LIABILITY, OR TORT (INCLUDING NEGLIGENCE OR OTHERWISE) ARISING IN ANY WAY OUT OF THE USE OF THIS SOFTWARE, EVEN IF ADVISED OF THE POSSIBILITY OF SUCH DAMAGE.

The other header files in `clang/lib/Headers/` are from clang with this license (see individual files for copyright owner):

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the "Software"), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

When using the ClamAV bytecode compiler to compile your own bytecode programs, you can release it under the license of your choice, provided that you comply with the license of the above header files.

APPENDIX A

Predefined macros

```
1 #define __llvm__ 1
2 #define __clang__ 1
3 #define __GNUC_MINOR__ 2
4 #define __GNUC_PATCHLEVEL__ 1
5 #define __GNUC__ 4
6 #define __GXX_ABI_VERSION 1002
7 #define __VERSION__ "4.2.1 Compatible Clang Compiler"
8 #define __STDC__ 1
9 #define __STDC_VERSION__ 199901L
10 #define __STDC_HOSTED__ 0
11 #define __CONSTANT_CFSTRINGS__ 1
12 #define __CHAR_BIT__ 8
13 #define __SCHAR_MAX__ 127
14 #define __SHRT_MAX__ 32767
15 #define __INT_MAX__ 2147483647
16 #define __LONG_MAX__ 9223372036854775807L
17 #define __LONG_LONG_MAX__ 9223372036854775807LL
18 #define __WCHAR_MAX__ 2147483647
19 #define __INTMAX_MAX__ 9223372036854775807L
20 #define __INTMAX_TYPE__ long int
21 #define __UINTMAX_TYPE__ long unsigned int
22 #define __INTMAX_WIDTH__ 64
23 #define __PTRDIFF_TYPE__ int
24 #define __PTRDIFF_WIDTH__ 32
25 #define __INTPTR_TYPE__ long int
26 #define __INTPTR_WIDTH__ 64
27 #define __SIZE_TYPE__ unsigned int
28 #define __SIZE_WIDTH__ 32
29 #define __WCHAR_TYPE__ int
30 #define __WCHAR_WIDTH__ 32
31 #define __WINT_TYPE__ int
32 #define __WINT_WIDTH__ 32
33 #define __SIG_ATOMIC_WIDTH__ 32
34 #define __FLT_DENORM_MIN__ 1.40129846e-45F
35 #define __FLT_HAS_DENORM__ 1
36 #define __FLT_DIG__ 6
37 #define __FLT_EPSILON__ 1.19209290e-7F
38 #define __FLT_HAS_INFINITY__ 1
39 #define __FLT_HAS_QUIET_NAN__ 1
40 #define __FLT_MANT_DIG__ 24
41 #define __FLT_MAX_10_EXP__ 38
42 #define __FLT_MAX_EXP__ 128
43 #define __FLT_MAX__ 3.40282347e+38F
44 #define __FLT_MIN_10_EXP__ (-37)
45 #define __FLT_MIN_EXP__ (-125)
46 #define __FLT_MIN__ 1.17549435e-38F
47 #define __DBL_DENORM_MIN__ 4.9406564584124654e-324
48 #define __DBL_HAS_DENORM__ 1
49 #define __DBL_DIG__ 15
50 #define __DBL_EPSILON__ 2.2204460492503131e-16
51 #define __DBL_HAS_INFINITY__ 1
52 #define __DBL_HAS_QUIET_NAN__ 1
53 #define __DBL_MANT_DIG__ 53
54 #define __DBL_MAX_10_EXP__ 308
55 #define __DBL_MAX_EXP__ 1024
56 #define __DBL_MAX__ 1.7976931348623157e+308
57 #define __DBL_MIN_10_EXP__ (-307)
58 #define __DBL_MIN_EXP__ (-1021)
59 #define __DBL_MIN__ 2.2250738585072014e-308
60 #define __LDBL_DENORM_MIN__ 4.9406564584124654e-324
61 #define __LDBL_HAS_DENORM__ 1
62 #define __LDBL_DIG__ 15
63 #define __LDBL_EPSILON__ 2.2204460492503131e-16
64 #define __LDBL_HAS_INFINITY__ 1
65 #define __LDBL_HAS_QUIET_NAN__ 1
66 #define __LDBL_MANT_DIG__ 53
67 #define __LDBL_MAX_10_EXP__ 308
68 #define __LDBL_MAX_EXP__ 1024
69 #define __LDBL_MAX__ 1.7976931348623157e+308
70 #define __LDBL_MIN_10_EXP__ (-307)
71 #define __LDBL_MIN_EXP__ (-1021)
72 #define __LDBL_MIN__ 2.2250738585072014e-308
73 #define __POINTER_WIDTH__ 64
```

```

#define __INT8_TYPE__ char
75 #define __INT16_TYPE__ short
#define __INT32_TYPE__ int
77 #define __INT64_TYPE__ long int
#define __INT64_C_SUFFIX__ L
79 #define __USER_LABEL_PREFIX__ _
#define __FINITE_MATH_ONLY__ 0
81 #define __GNUC_STDC_INLINE__ 1
#define __NO_INLINE__ 1
83 #define __FLT_EVAL_METHOD__ 0
#define __FLT_RADIX__ 2
85 #define __DECIMAL_DIG__ 17
#define __CLAMBC__ 1
87 #define __BYTECODE_API_H
#define __EXECS_H
89 #define __BC_FEATURES_H
#define __EBOUNDS(x)
91 #define __PE_H
#define __DISASM_BC_H
93 #define __BYTECODE_DETECT_H
#define __STDBOOL_H
95 #define __bool__ bool
#define __true__ true
97 #define __false__ false
#define __bool_true_false_are_defined__ 1
99 #define __force_inline__ inline __attribute__((always_inline))
#define __VIRUSNAME_PREFIX(name) const char __clambc_virusname_prefix[] = name;
101 #define __VIRUSNAMES(...) const char *const __clambc_virusnames[] = {__VA_ARGS__};
#define __PE_UNPACKER_DECLARE const uint16_t __clambc_kind = BC_PE_UNPACKER;
103 #define __SIGNATURES_DECL_BEGIN struct __Signatures {
#define __DECLARE_SIGNATURE(name) const char *name##_sig; __Signature name;
105 #define __SIGNATURES_DECL_END };
#define __TARGET(tgt) const unsigned short __Target = (tgt);
107 #define __COPYRIGHT(c) const char *const __Copyright = (c);
#define __ICONGROUP1(group) const char *const __IconGroup1 = (group);
109 #define __ICONGROUP2(group) const char *const __IconGroup2 = (group);
#define __FUNCTIONALITY_LEVEL_MIN(m) const unsigned short __FuncMin = (m);
111 #define __FUNCTIONALITY_LEVEL_MAX(m) const unsigned short __FuncMax = (m);
#define __SIGNATURES_DEF_BEGIN static const unsigned __signature_bias =
__COUNTER__ + 1; const struct __Signatures Signatures = {
113 #define __DEFINE_SIGNATURE(name, hex) .name##_sig = (hex), .name = {__COUNTER__ -
__signature_bias},
#define __SIGNATURES_END };
115 #define __RE2C_BSIZE 128
#define __YYCTYPE unsigned char
117 #define __YYCURSOR re2c_scur
#define __YYLIMIT re2c_slm
119 #define __YYMARKER re2c_smrk
#define __YYCONTEXT re2c_sctx
121 #define __YYFILL(n) { RE2C_FILLBUFFER(n); if (re2c_sres <= 0) break;}
#define __REGEX_SCANNER unsigned char *re2c_scur, *re2c_stok, *re2c_smrk, *re2c_sctx,
*re2c_slm; int re2c_sres; int32_t re2c_stokstart; unsigned char
re2c_sbuffer[RE2C_BSIZE]; re2c_scur = re2c_slm = re2c_smrk = re2c_sctx =
&re2c_sbuffer[0]; re2c_sres = 0; RE2C_FILLBUFFER(0);
123 #define __REGEX_POS -(re2c_slm - re2c_scur) + seek(0, SEEK_CUR))
#define __REGEX_LOOP_BEGIN do { re2c_stok = re2c_scur; re2c_stokstart = REGEX_POS;}
while (0);
125 #define __REGEX_RESULT (re2c_sres)
#define __RE2C_DEBUG_PRINT do { char buf[81]; uint32_t here = seek(0, SEEK_CUR);
uint32_t d = re2c_slm - re2c_scur; uint32_t end = here - d; unsigned len =
end - re2c_stokstart; if (len > 80) { unsigned skipped = len - 74;
seek(re2c_stokstart, SEEK_SET); if (read(buf, 37) == 37) break; memcpy(buf+37,
"[...]", 5); seek(end-37, SEEK_SET); if (read(buf, 37) != 37) break; buf[80] =
'\0'; } else { seek(re2c_stokstart, SEEK_SET); if (read(buf, len) != len)
break; buf[len] = '\0'; } buf[80] = '\0'; debug_print_str(buf, 0); seek(here,
SEEK_SET);} while (0)
127 #define __DEBUG_PRINT_REGEX_MATCH RE2C_DEBUG_PRINT
#define __BUFFER_FILL(buf, cursor, need, limit) do { (limit) = fill_buffer((buf),
sizeof(buf), (limit), (cursor), (need));} while (0);
129 #define __BUFFER_ENSURE(buf, cursor, need, limit) do { if ((cursor) + (need) >= (limit))
{ BUFFER_FILL(buf, cursor, need, limit) (cursor) = 0; } } while (0);
#define __RE2C_FILLBUFFER(need) do { uint32_t cursor = re2c_stok - &re2c_sbuffer[0];
int32_t limit = re2c_slm - &re2c_sbuffer[0]; limit =
fill_buffer(re2c_sbuffer, sizeof(re2c_sbuffer), limit, (cursor), (need)); if
(!limit) { re2c_sres = 0; } else if (limit <= (need)) { re2c_sres = -1; } else
{ uint32_t curoff = re2c_scur - re2c_stok; uint32_t mrkoff = re2c_smrk -
re2c_stok; uint32_t ctxoff = re2c_sctx - re2c_stok; re2c_slm =
&re2c_sbuffer[0] + limit; re2c_stok = &re2c_sbuffer[0]; re2c_scur =
&re2c_sbuffer[0] + curoff; re2c_smrk = &re2c_sbuffer[0] + mrkoff; re2c_sctx =
&re2c_sbuffer[0] + ctxoff; re2c_sres = limit; } } while (0);

```