

Introductory Firmware Reverse Engineering

Canon Selphi printer

Aleksandar Nikolic, BSides PDX 2025

Tools to install

- Ghidra
 - [https://github.com/NationalSecurityAgency/ghidra/releases/tag/Ghidra 11.4.2 build](https://github.com/NationalSecurityAgency/ghidra/releases/tag/Ghidra_11.4.2_build)
 - May need JRE as well
- Binwalk
 - Probably through your distribution
- ImHex
 - <https://imhex.werwolv.net/>

Workshop materials

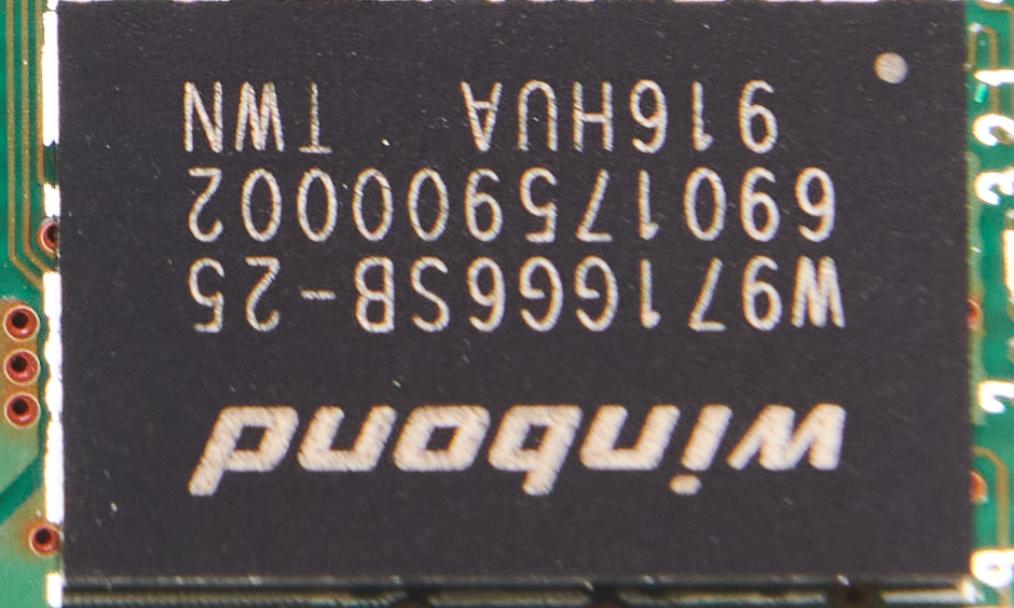
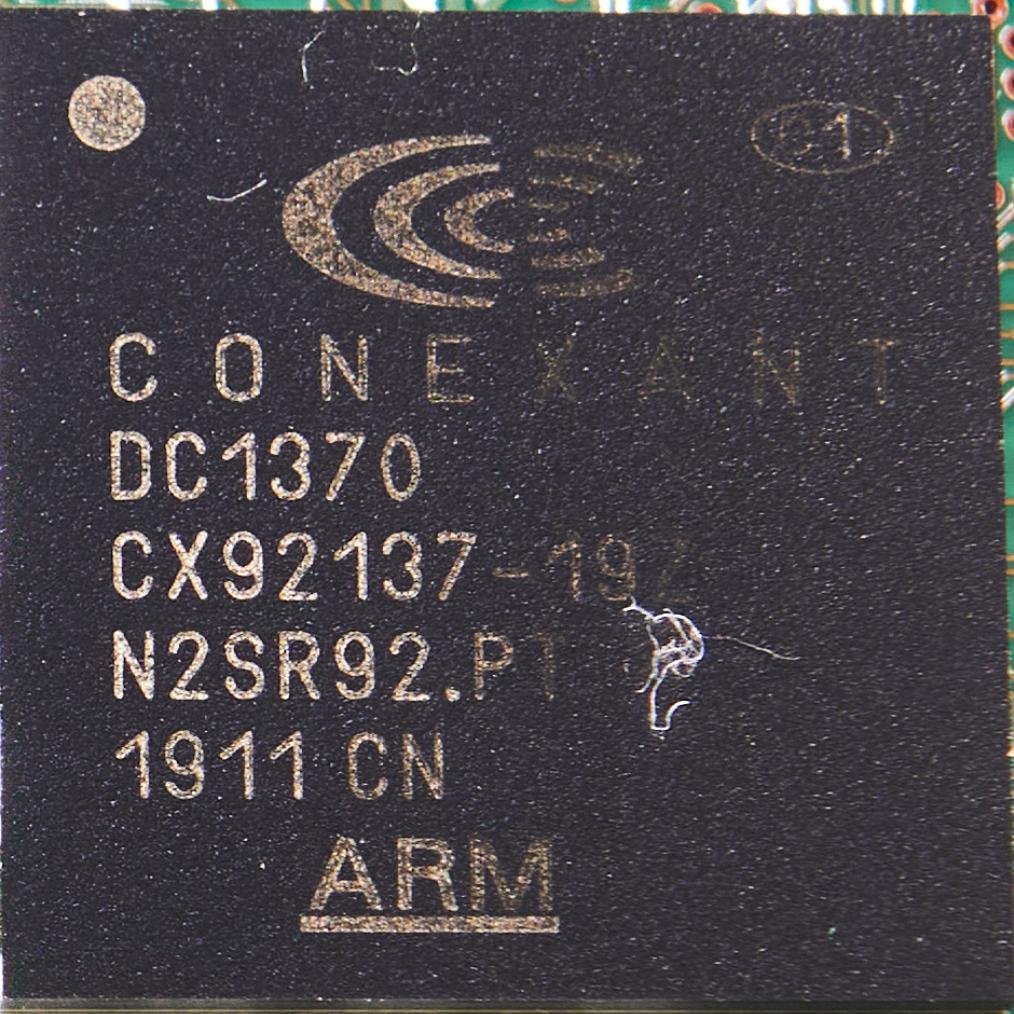
- [https://github.com/ea/
selphy workshop](https://github.com/ea/selphy_workshop)
-

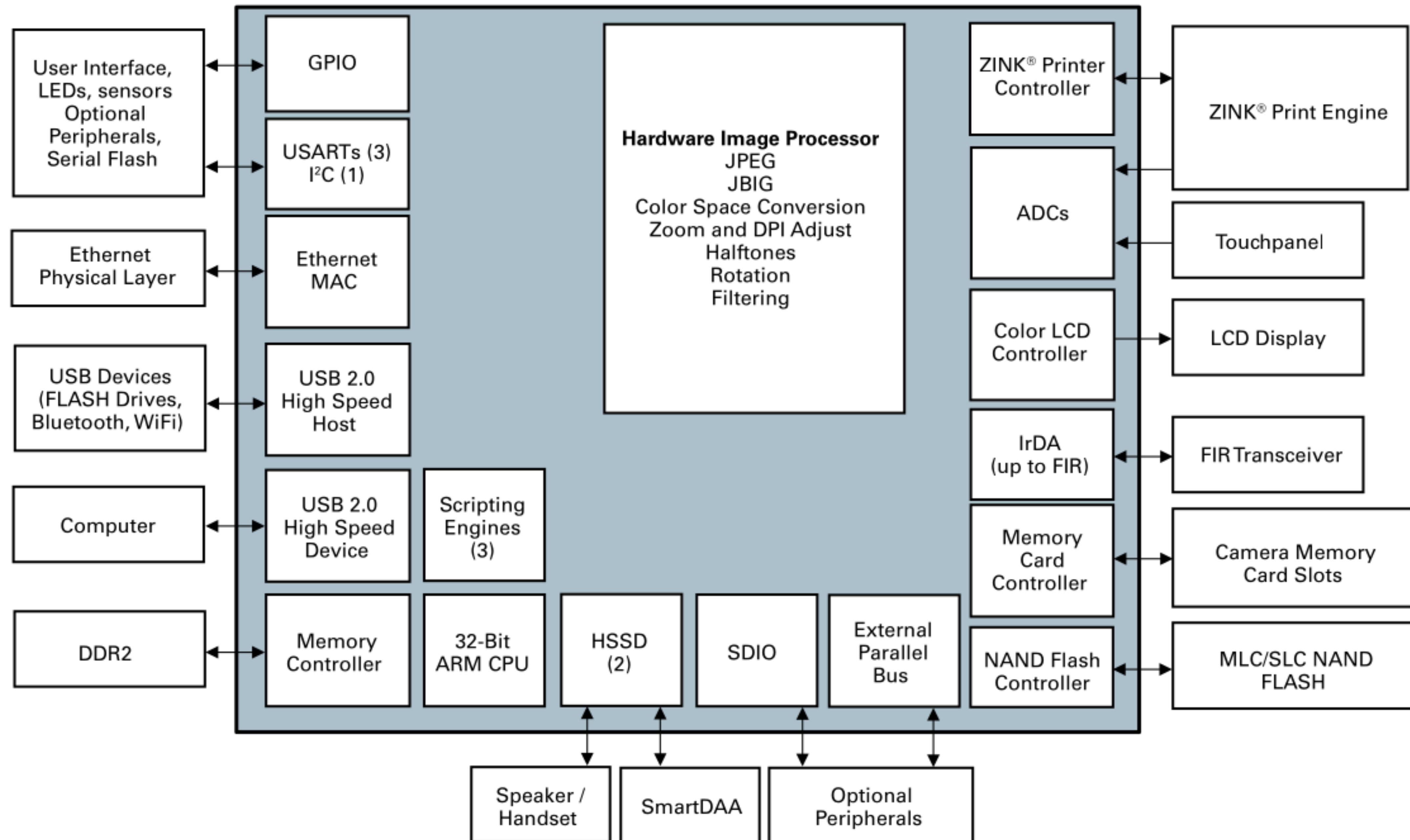




2EE-01536AD

007-AE0325
MODEL: WM320
FCC ID: AZD320
IC: 498J-320
CANON INC.

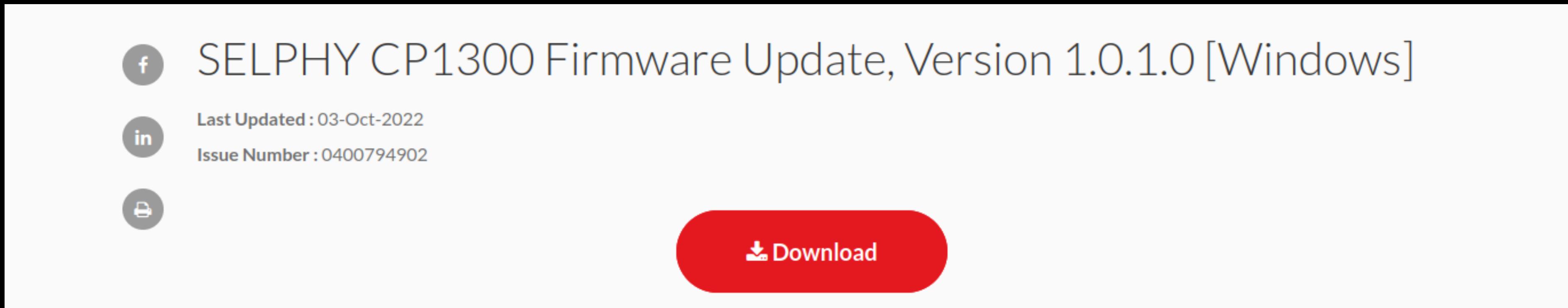




CX92137 Block Diagram

Step 0: Dive right in

- Grab the firmware
- Run it through binwalk
- <https://my.canon/en/support/0400794902>
-



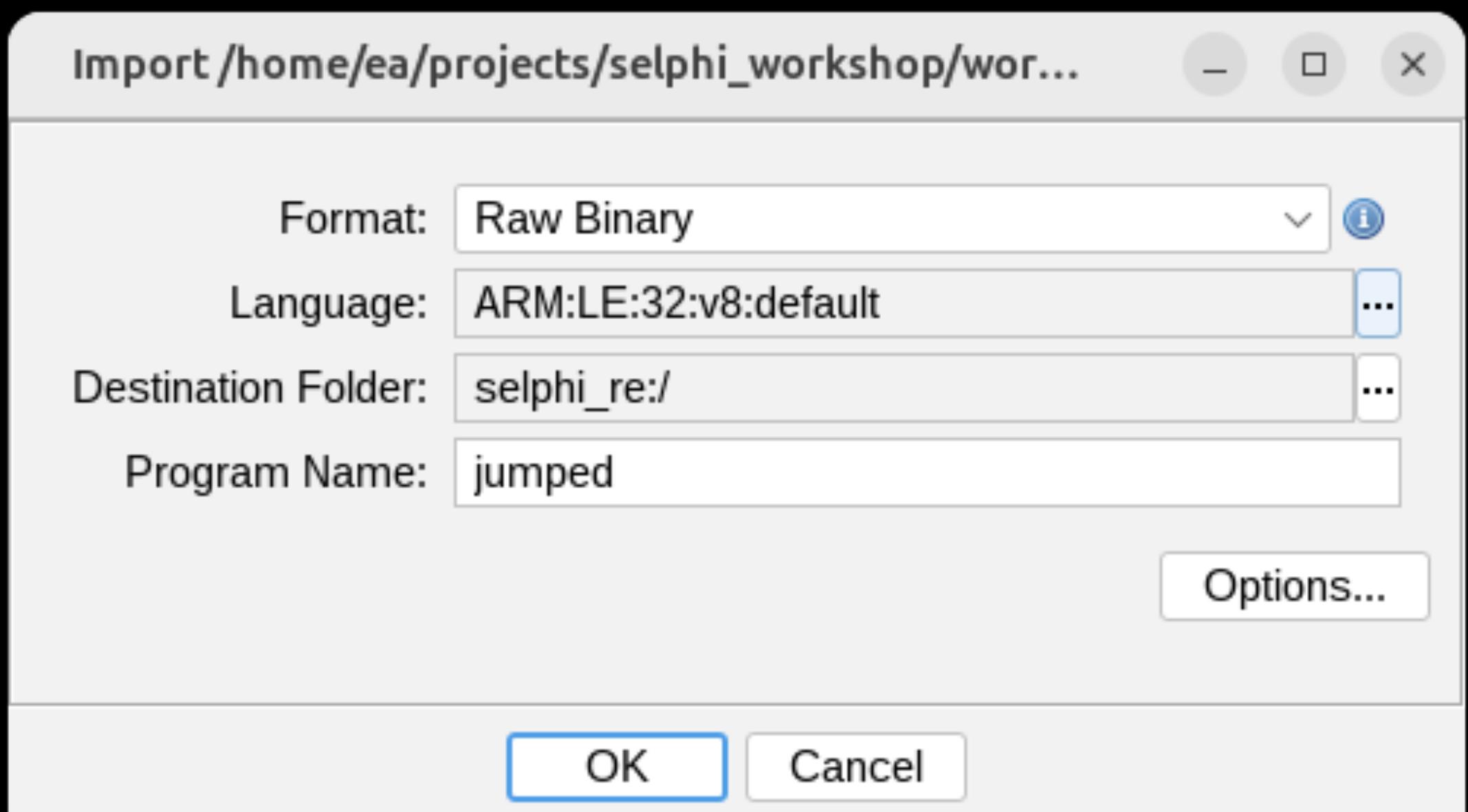
Binwalk

- Recognize known file types
- Recursively extract known archive file formats
- binwalk -e

DECIMAL	HEXADECIMAL	DESCRIPTION
183886	0x2CE4E	Copyright string: "Copyright (c) 1996-201
656892	0xA05FC	gzip compressed data, from NTFS filesystem
3941358	0x3C23EE	MySQL ISAM index file Version 7
3946106	0x3C367A	MySQL MISAM index file Version 7
3960252	0x3C6DBC	MySQL MISAM index file Version 2
3960455	0x3C6E87	MySQL MISAM index file Version 2

Step 1: Load into Ghidra

- We'll talk about details later...



Key terms: Disassembly

- Turning bytes into readable assembly code
- Non-trivial – often imperfect
- We don't know what's data and what's code

00089f5c d5 0f 4f e2	adr	r0,DAT_00089c10
00089f60 8f fd fd eb	bl	printf?
00089f64 00 40 8d e5	str	r4,[sp,#0x0]=>local_18
00089f68 0c 00 95 e5	ldr	r0,[r5,#0xc]
00089f6c 07 30 a0 e1	cpy	r3,r7
00089f70 05 20 a0 e1	cpy	r2,r5
00089f74 06 10 a0 e1	cpy	r1,r6
00089f78 8d fd ff eb	bl	FUN_000895b4
00089f7c 02 00 d4 e5	ldrb	r0,[r4,#0x2]
00089f80 00 00 50 e3	cmp	r0,#0x0
00089f84 07 00 00 0a	beq	LAB_00089fa8
00089f88 1c 03 9f e5	ldr	r0,[DAT_0008a2ac]
00089f8c 02 10 84 e2	add	r1,r4,#0x2
00089f90 d8 66 0f eb	bl	FUN_00463af8
00089f94 10 13 9f e5	ldr	r1,[DAT_0008a2ac]
00089f98 31 0e 8f e2	adr	r0,s_dpof_file_path[%s]_0008a2b0
00089f9c 80 fd fd eb	bl	printf?
00089fa0 00 00 a0 e3	mov	r0,#0x0
XREF[1]:		
00089fa4 f8 80 bd e8	ldmia	sp!,{r3,r4,r5,r6,r7,pc}
XREF[1]:		
00089fa8 ff 00 a0 e3	mov	r0,#0xff
00089fac fc ff ff ea	b	LAB_00089fa4
00089fb0 e4	??	E4h
00089fb1 02	??	02h
00089fb2 9f	??	9Fh
00089fb3 e5	??	E5h
00089fb4 0c	??	0Ch
00089fb5 00	??	00h
00089fb6 90	??	90h
00089fb7 e5	??	E5h
00089fb8 1e	??	1Eh
00089fb9 ff	??	FFh
00089fba 2f	??	2Fh /

Key terms: Decompiling

```
main(int argc, char **argv){  
    int sum_of_even = 0;  
    int sum_of_odd = 0;  
    for(int i = 0; i < atoi(argv[1]); i++){  
        if(i % 2 == 0) sum_of_even += i;  
        else sum_of_odd +=1;  
    }  
    printf("even %d, odd %d\n"  
          ,sum_of_even, sum_of_odd);  
}
```

		LAB_0040188f	
0040188f	8b 45 fc	MOV	EAX,dword ptr [RBP + -0x4]
00401892	83 e0 01	AND	EAX,0x1
00401895	85 c0	TEST	EAX,EAX
00401897	75 08	JNZ	LAB_004018a1
00401899	8b 45 fc	MOV	EAX,dword ptr [RBP + -0x4]
0040189c	01 45 f4	ADD	dword ptr [RBP + -0xc],EAX
0040189f	eb 04	JMP	LAB_004018a5
		LAB_004018a1	
004018a1	83 45 f8 01	ADD	dword ptr [RBP + -0x8],0x1
		LAB_004018a5	
004018a5	83 45 fc 01	ADD	dword ptr [RBP + -0x4],0x1
		LAB_004018a9	
004018a9	48 8b 45 e0	MOV	RAX,qword ptr [RBP + -0x20]
004018ad	48 83 c0 08	ADD	RAX,0x8
004018b1	48 8b 00	MOV	RAX,qword ptr [RAX]
004018b4	48 89 c7	MOV	RDI,RAX
004018b7	b8 00 00 00 00	MOV	EAX,0x0
004018bc	e8 4f 2c	CALL	FUN_00404510

```
undefined8 FUN_00401865(undefined8 param_1, long param_2)  
{  
    int iVar1;  
    int local_14;  
    int local_10;  
    uint local_c;  
  
    local_14 = 0;  
    local_10 = 0;  
    local_c = 0;  
    while( true ) {  
        iVar1 = FUN_00404510(*(undefined8 *)(&param_2 + 8));  
        if (iVar1 <= (int)local_c) break;  
        if ((local_c & 1) == 0) {  
            local_14 = local_14 + local_c;  
        }  
        else {  
            local_10 = local_10 + 1;  
        }  
        local_c = local_c + 1;  
    }  
    FUN_00405770("even %d, odd %d\n",local_14,local_10);  
    return 0;  
}
```

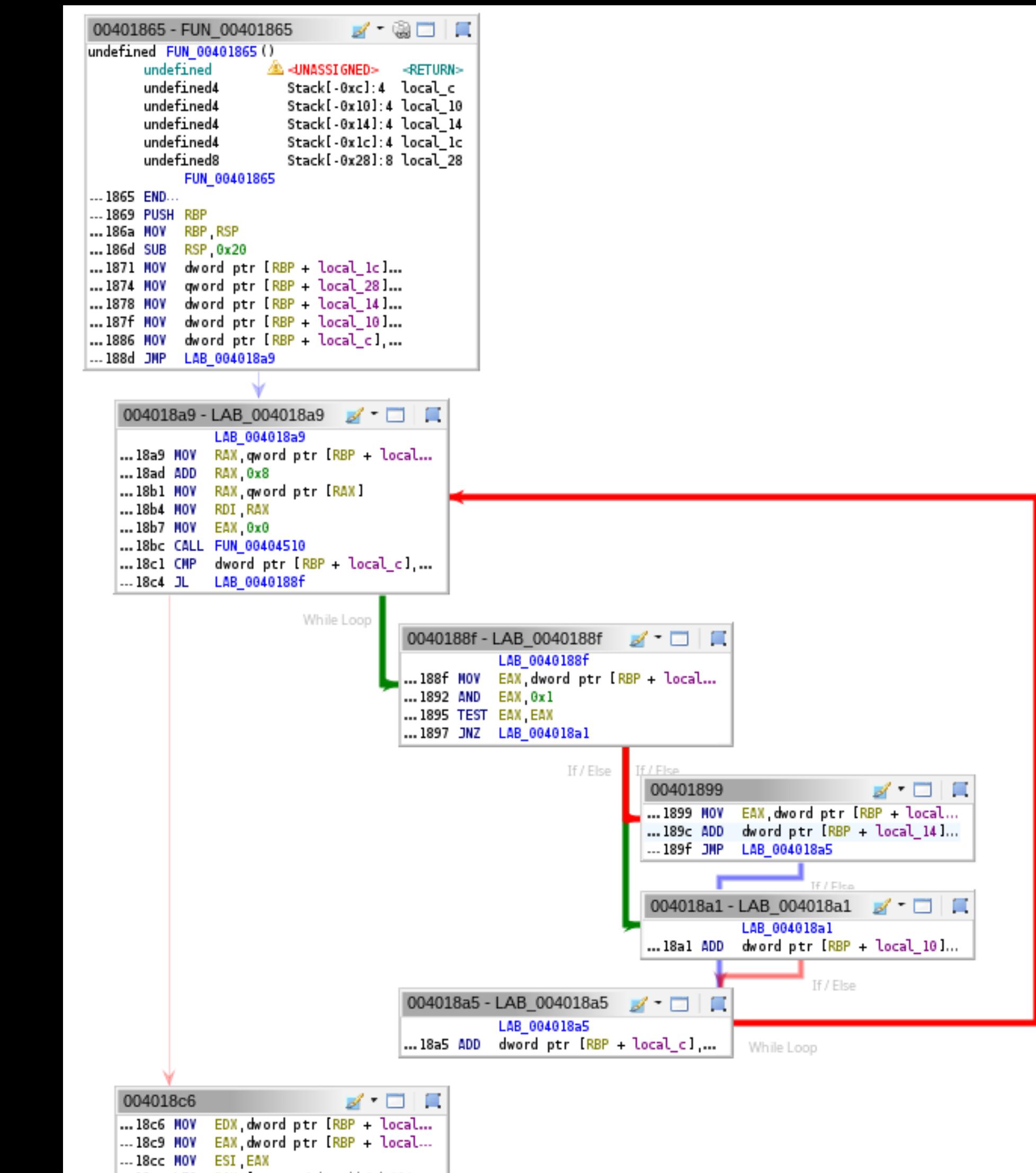
Key terms: Symbols

- Function names
- Variable names
- Type information
 - Type names, structure definitions, sizes...
- With firmware , we usually don't have them

```
const char *callReplyGetBigNumber(CallReply *rep, size_t *len) {
    callReplyParse(rep);
    if (rep->type != REDISMODULE_REPLY_BIG_NUMBER) return NULL;
    *len = rep->len;
    return rep->val.str;
}
```

Key terms: basic blocks

- Functions are graphs of basic blocks
- Jump instructions form edges



Hopefully, the analysis is done.

Ghidra UI

- Config and overview
- Emphasis on decompiler & XREFS

Key concepts: XREFS

- How we navigate through the binary
- Think in graphs!
- X has a relationship with Y
- Function A calls function B
- Instruction I uses memory address C
-
- Without reference resolution and propagation, a tool is useless

XREFs in Ghidra

- Not visible in decompiler

```
*****  
*  
*          FUNCTION  
*****  
  
undefined undefined FUN_00023240()  
⚠<UNASSIGNED>  <RETURN>  
undefined4 Stack[-0x1c]:4 local_1c  
undefined4 Stack[-0x24]:4 local_24  
undefined4 Stack[-0x28]:4 local_28  
  
undefined4 Stack[-0x2c]:4 local_2c  
undefined4 Stack[-0x30]:4 local_30  
undefined4 Stack[-0x34]:4 local_34  
undefined4 Stack[-0x38]:4 local_38  
  
FUN_00023240  
  
XREF[1]: 000232f4(W)  
XREF[1]: 000232f8(*)  
XREF[1,1]: 000232bc(W),  
           000232d0(R)  
  
XREF[1]: 000232b8(W)  
XREF[1]: 000232b4(W)  
XREF[1]: 00023270(W)  
XREF[2]: 000232ac(W),  
           000232fc(R)  
  
XREF[25]: FUN_00003cf:00003d18(c),  
           FUN_0000449c:00004584(c),  
           FUN_0000449c:000045a0(c),  
           dc_rom_resource_setup:000046e0(c...  
           FUN_00007bf8:00007c78(c),  
           FUN_00009fdc:0000a010(c),  
           FUN_0000a018:0000a06c(c),  
           FUN_000235ac:00023690(c),  
           FUN_000279e0:00027a00(c),  
           FUN_0002e088:0002e0bc(c),  
           FUN_0002e088:0002e0f8(c),  
           FUN_0002e088:0002e134(c),  
           FUN_0002e088:0002e170(c),  
           FUN_0004df30:0004df50(c),  
           FUN_0004df30:0004df6c(c),  
           FUN_0004df30:0004df88(c),  
           FUN_0004df30:0004dfa4(c),  
           FUN_0004df30:0004dfc0(c),  
           FUN_0004df30:0004dfdc(c),  
           FUN_0004e594:0004e5b4(c), [more]  
  
00023240 f0 41 2d e9      stmdb      sp!, {r4, r5, r6, r7, r8, lr}  
00023241 01 00 00 01
```

Recovering context

- Easy clues
 - Symbols?
 - Strings?
 - Well known functions
- Harder
 - Datasheets, architecture manuals

Naming things

- Adopt a convention and stick to it
- Function names – guess but append “?” at the end
- Variables – give hints about type – “psMsg”
-
- Keep guessing, refine as you get more info

Key concept: Structures

- C Structs are everywhere
- Data accessed at offset into structure
- Crucial for references
- All languages map to them

```
struct msg_hdr {  
    int size;  
    int offset;  
    int value;  
};
```

```
#include <stdio.h>

struct test_struct {
    int size;
    int offset;
    int value;
};

void print_struct(struct test_struct *a){
    printf("size %x, offset %x, value %d\n",
           a->size, a->offset, a->value);
}

int main(int argc, char **argv){
    struct test_struct a;
    a.size = 10;
    a.offset = 0xff;
    a.value = -4;
    print_struct(&a);
}
```

```
undefined8 FUN_001011a6(void)
{
    long in_FS_OFFSET;
    undefined4 local_1c;
    undefined4 local_18;
    undefined4 local_14;
    long local_10;

    local_10 = *(long *)(in_FS_OFFSET + 0x28);
    local_1c = 10;
    local_18 = 0xff;
    local_14 = 0xfffffffffc;
    FUN_00101169(&local_1c);
    if (local_10 != *(long *)(in_FS_OFFSET + 0x28)) {
        /* WARNING: Subroutine does not
         *          _stack_chk_fail();
    }
    return 0;
}
```

Structure Editor - struct_1 (structures) [CodeBrowser: structs:/structures]

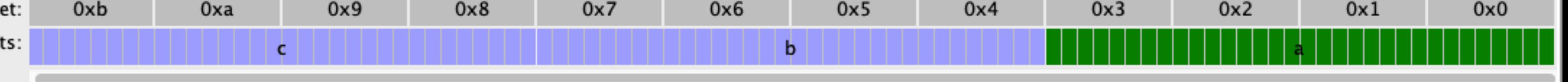
Edit Help

Structure Editor – struct_1 (structures)

Offset	Length	Mnemonic	DataType	Name	Comment
0x0	0x4	uint32_t	uint32_t	a	
0x4	0x4	uint32_t	uint32_t	b	
0x8	0x4	uint32_t	uint32_t	c	

Search:

Byte Offset: 0xb 0xa 0x9 0x8 0x7 0x6 0x5 0x4 0x3 0x2 0x1 0x0

Component Bits: 

Category: structures/

Name: struct_1

Description: c

Size: 0xc Alignment: 0x1

align (minimum)
 default

 machine: 2

pack
 default

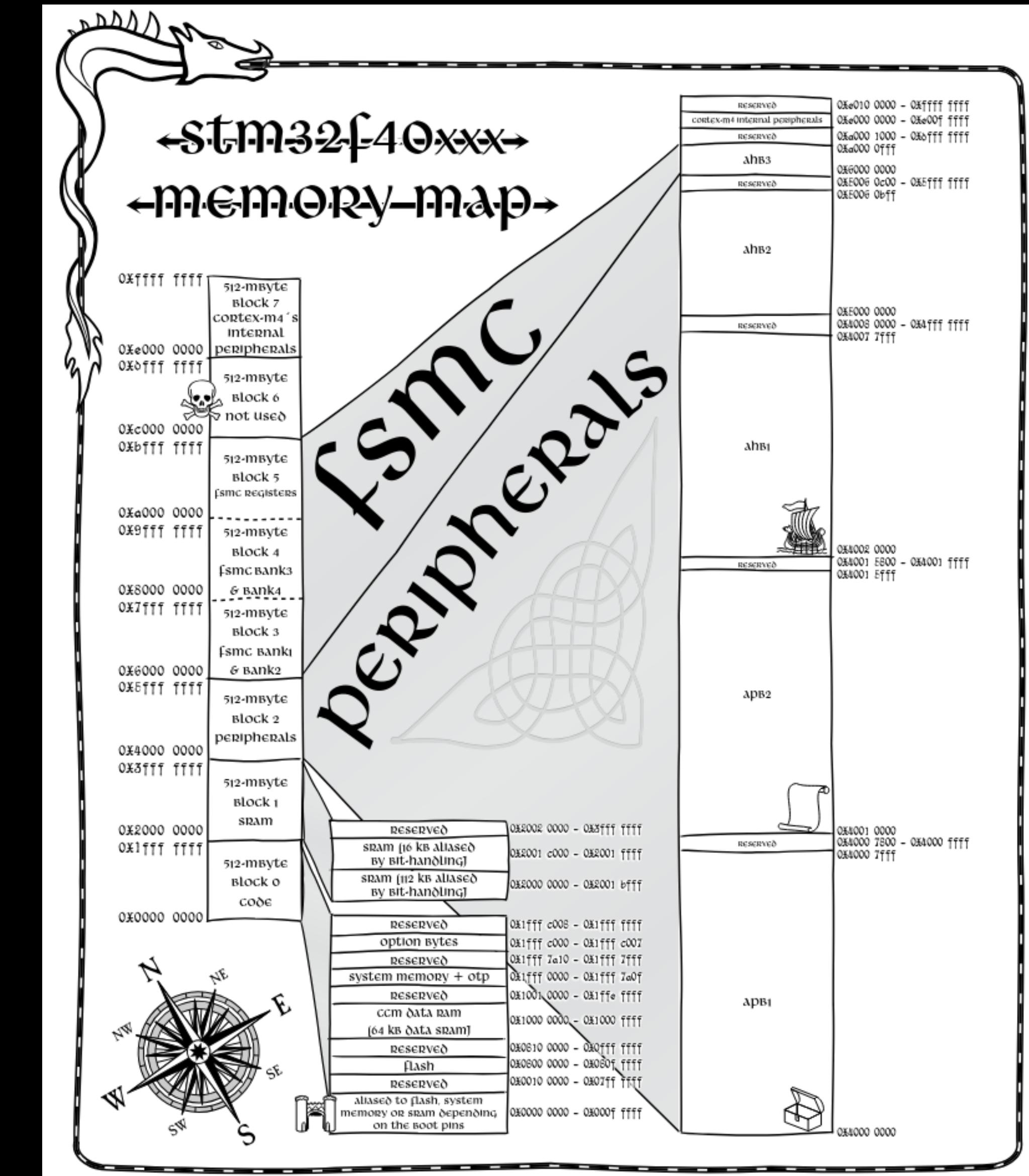
```
test local_1c;

lVar1 = *(long *)(in_FS_OFFSET + 0x28);
local_1c.a = 10;
local_1c.b = 0xff;
local_1c.c = 0xffffffff;
FUN_00101169(&local_1c);
```

Let's go digging through the FW.

Key concept: Memory map

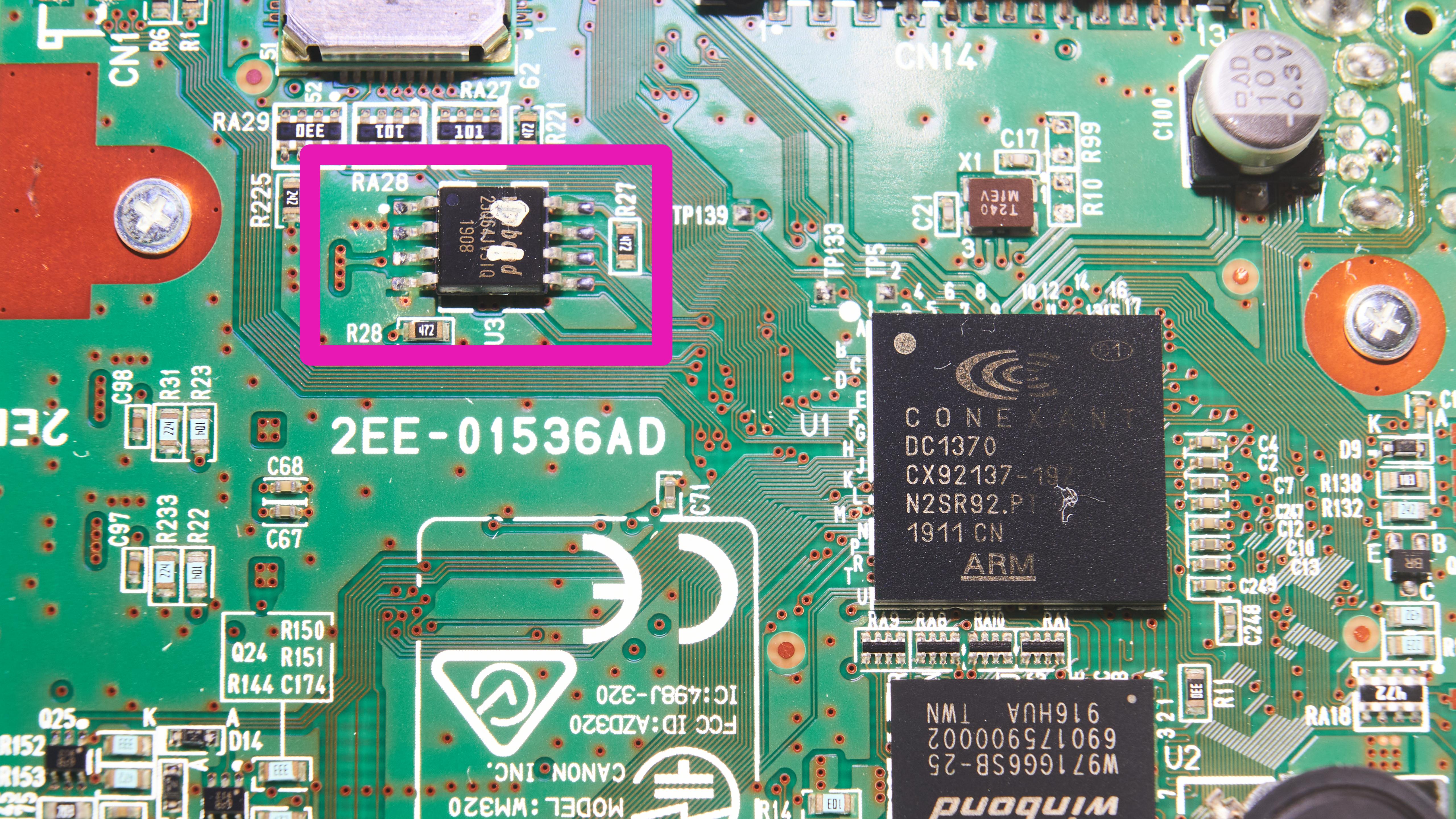
- Not everything starts at 0
- Firmware load address?
- Flash memory?
- Peripherals?
-
- Ideally you have it, we don't



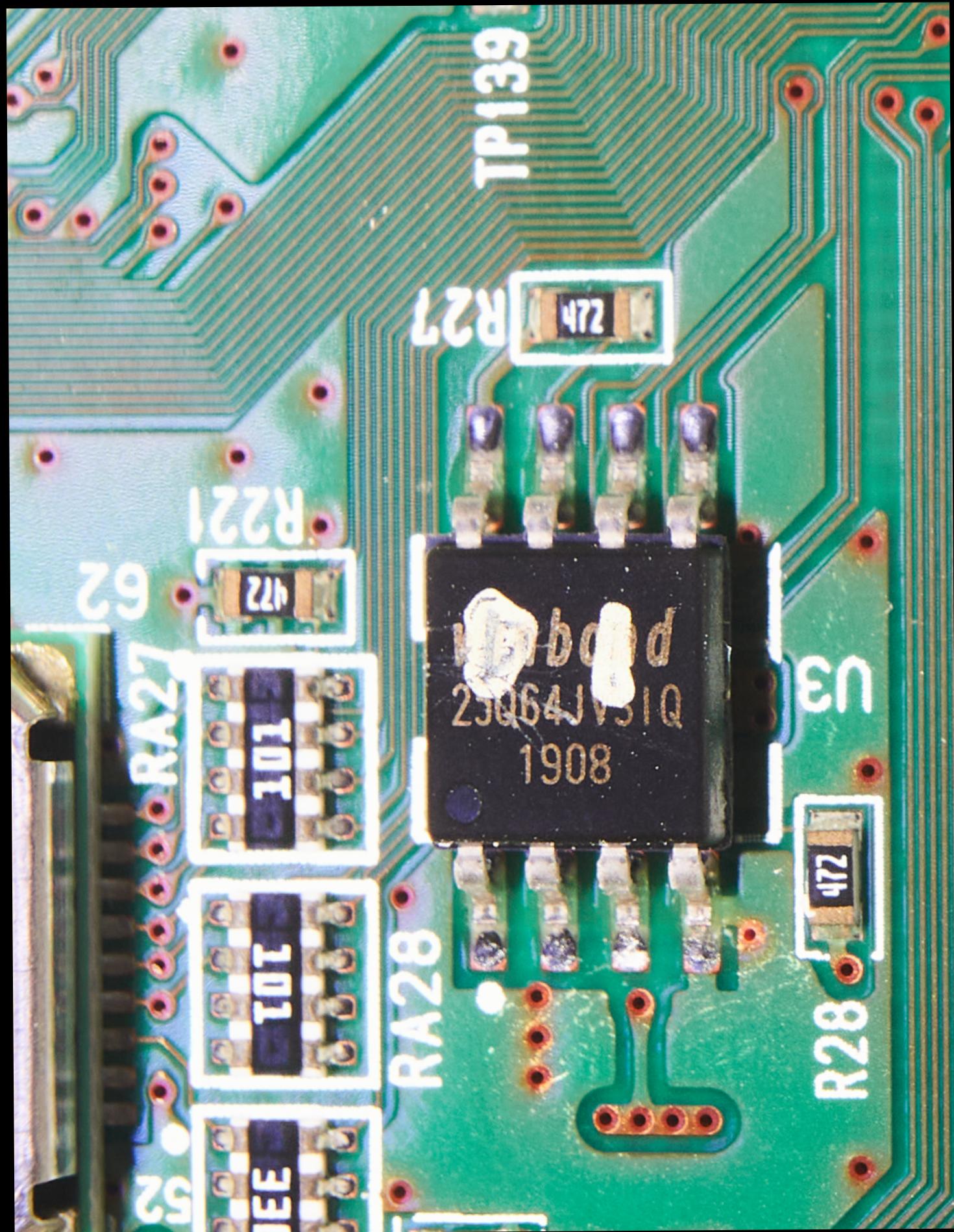
**How do we figure it out?
Firmware update might have a hint!**

Lets look at the better mapping

Let's talk about firmware
dumping.



SPI Flash



W25Q64JV



3. PACKAGE TYPES AND PIN CONFIGURATIONS

3.1 Pin Configuration SOIC 150/208-mil

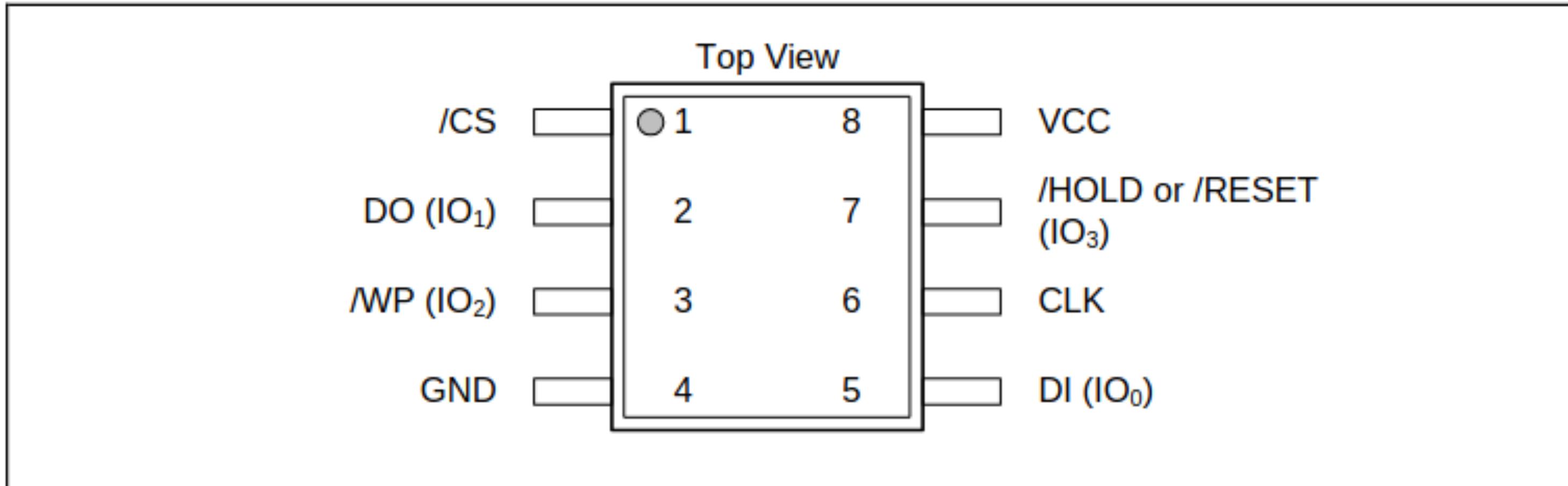


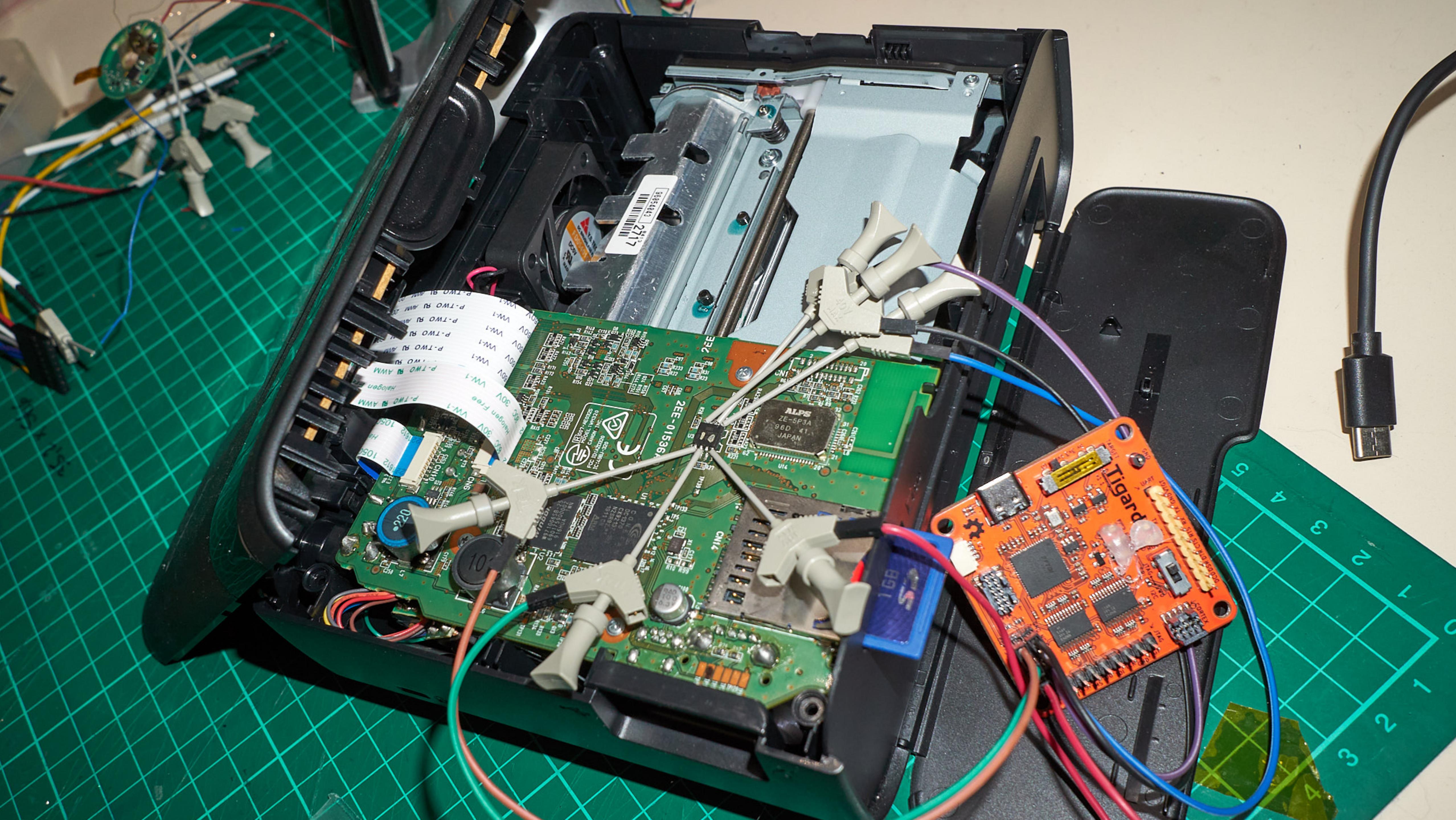
Figure 1a. W25Q64JV Pin Assignments, 8-pin SOIC 208-mil (Package Code SN,SS)

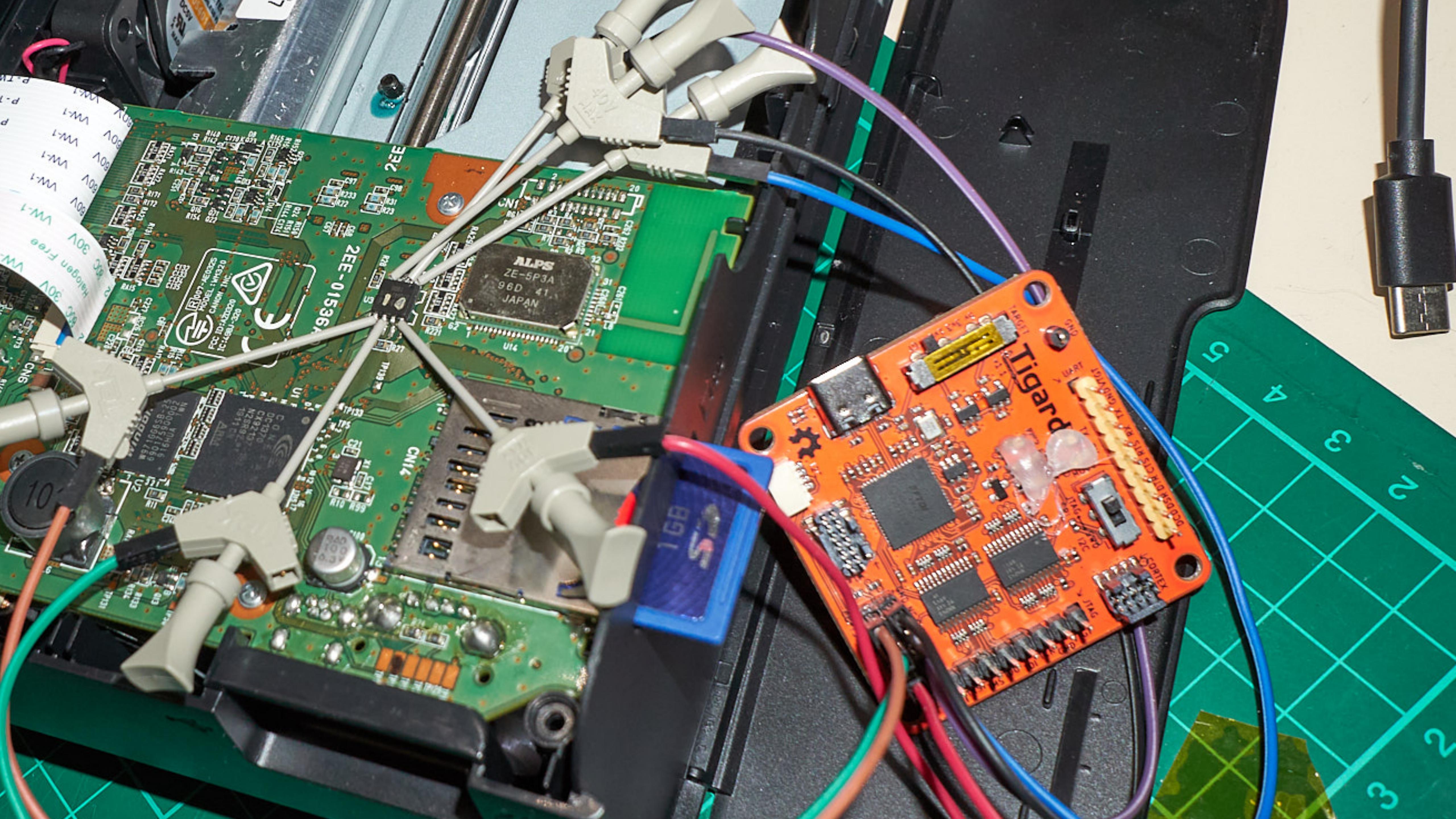
Reading SPI Flash

- SPI – serial peripheral interface
 - Relatively slow serial communication protocol
- Flashrom – implements reading most known chips
- Reader?
 - A Pi Pico will do
 - <https://github.com/stacksmashing/pico-serprog>
 - Or a Tigard
 - <https://www.crowdsupply.com/securinghw/tigard>
 -

 flashrom

The logo for flashrom, featuring the word "flashrom" in a blue, lowercase, sans-serif font. The letter "a" has a yellow lightning bolt icon through its center.





Flash dump
Let's explore the diff.

**Firmware update
There's always a checksum.**

```
uVar1 = thunk_FUN_06f6e614(param_1,0x7f4e18,uVar4);
if (uVar1 != uVar4) {
    FUN_06ed5040();
    fmt = (char *)0x38ad0;
    goto LAB_06ed1bb4;
}
for (uVar1 = 0; uVar1 < uVar4; uVar1 = uVar1 + 4) {
    uVar2 = uVar2 ^ *(uint *) (uVar1 + 0x7f4e18);
}
uVar3 = uVar3 + uVar4;
} while( true );
```

```
romtab:
    magic: 1414676815
    rom_version: 0x15430002
    unk1: 0xa
    num_items: 52
    fw_idx: 0x1c
    fw_memory_address: 0x6e98000
    romtab_size: 5
    rom_size: 0x567d8c
    unk5: 0xfc036ae
    unk6: 0x32db
    unk7: 0x80000000
```

```
def get_checksum(filename, checksum_offset, romtab_offset):
    with open(filename, 'rb') as f:
        f.seek(checksum_offset)
        byte_block = f.read(4)
        orig_checksum = int.from_bytes(byte_block, 'little')
        f.seek(romtab_offset)
        result = 0
        for i in range(0, os.path.getsize(f.name), 4):
            byte_block = f.read(4)
            dword = int.from_bytes(byte_block, 'little')
            result ^= dword
        if result != 0:
            print("Correct checksum should be: %x %%(%result^orig_checksum))")
    return result^orig_checksum
```

Update the firmware?
1.0.2._->1.0.2._

OK

Cancel



Busy



So that's it?

- There's one more tricky bit:

```
$ file FIRMWARE
FIRMWARE: gzip compressed data, last modified: Tue Jan  4 10:58:18 2022, from NTFS filesystem
$ ls -lh FIRMWARE
-rw-rw-r-- 1 ea ea 3.9M Sep  7 20:15 FIRMWARE
$ cp FIRMWARE FIRMWARE_unpack.gz
$ gunzip FIRMWARE_unpack.gz
$ ls -lh FIRMWARE_unpack
-rw-rw-r-- 1 ea ea 7.5M Sep  7 20:16 FIRMWARE_unpack
$ gzip FIRMWARE_unpack
$ ls -lh FIRMWARE_unpack.gz
-rw-rw-r-- 1 ea ea 3.5M Sep  7 20:16 FIRMWARE_unpack.gz
```

```
    }
    uVar3 = get_bits(&fw_update, 0x10);
    *(undefined2 *)&fw_update->field_0xa8 = uVar3;
    FUN_000023cc(fw_update, 0x10);
    FUN_00002420(fw_update, 0x10);
    uVar2 = 1;
}
LAB_000020cc:
    fw_update->field136_0xa0 = uVar2;
}
else {
    if (cVar1 != '\x01') {
        if (cVar1 != '\x02') {
            printf("INFLATE() error");
            do {
                /* WARNING: Do nothing block with infinite loop */
                } while( true );
            }
        FUN_00002478(fw_update);
        uVar2 = 0;
        goto LAB_000020cc;
    }
    uVar4 = (uint)*(ushort *)&fw_update->field_0xa8;
    fw_update->field136_0xa0 = uVar4 != 0;
    if ((int)*(uint *)&fw_update->expected_size < (int)uVar4) {
        uVar4 = *(uint *)&fw_update->expected_size;
    }
    btype = fw_update->field3_0xc;
    uVar5 = uVar4 & 0xffff;
    fw_update->field3_0xc = btype + uVar5;
    *(uint *)&fw_update->expected_size = *(int *)&fw_update->expected_size - uVar5;
    *(short *)&fw_update->field_0xa8 = *(short *)&fw_update->field_0xa8 - (short)uVar4;
    pcVar6 = fw_update->load_address + btype;
    while (bVar7 = uVar5 != 0, uVar5 = uVar5 - 1 & 0xffff, bVar7) {
        cVar1 = FUN_00001fd0(fw_update->archive_data);
        *pcVar6 = cVar1;
        fw_update->archive_data = fw_update->archive_data + 1;
        pcVar6 = pcVar6 + 1;
    }
}
```



DEFLATE algorithm

- First bit: Last-block-in-stream marker:
 - `1`: This is the last block in the stream.
 - `0`: There are more blocks to process after this one.
- Second and third bits: Encoding method used for this block type:
 - `00`: A stored (a.k.a. raw or literal) section, between 0 and 65,535 bytes in length
 - `01`: A *static Huffman* compressed block, using a pre-agreed Huffman tree defined in the RFC
 - `10`: A *dynamic Huffman* compressed block, complete with the Huffman table supplied
 - `11`: Reserved—don't use.

Infgen

from Mark Adler

- <https://github.com/madler/infgen>

```
$ ./infgen -h
infgen 3.5
Usage:
infgen [-d[d]misc[c]q[q]rb[b]] input_path > output_path
infgen [-d[d]misc[c]q[q]rb[b]] < input_path > output_path

-d  Write raw dynamic header (code lengths in comments)
-dd Also show the bits for each element displayed
-m  Show copied data after each match
-i  Include detailed gzip / zlib header descriptions
-s  Include deflate block statistics (as comments)
-c  Color the output components (if terminal supports)
-cc Use high-intensity instead of standard colors
-q  Do not write dynamic code lengths (comments or not)
-qq Do not write deflate stream description at all
-r  Assume raw deflate data -- do not look for headers
-b  Write compact binary format (only -r honored)
-bb Write compact binary format with bit counts
```

Original

```
$ ./infgen -bb < ../flash_dump/FIRMWARE | ./infstats
1 stream
31 stored blocks: 557,307(4) -> 557,153 (1.000)
    data: μ = 17,972.7, σ = 1,027.9, in [16,383..19,160]
136 fixed blocks: 3,470,468(6) -> 7,234,335 (0.480)
    lits: 1,491,990(2) -> 1,432,053 (1.042)
    mats: 1,978,308(4) -> 5,802,282 (0.341)
    bits: μ = 204,145.2, σ = 25,244.1, in [42,869..241,189]
    syms: μ = 16,286.6, σ = 1,124.1, in [3,274..16,383]
    lens: μ = 7.4, σ = 16.8, in [3..258]
    data: μ = 53,193.6, σ = 27,264.0, in [19,211..233,466]
```

Modified

```
$ ./infgen -bb < ../flash_dump/mod/FIRMWARE_test.gz | ./infstats
1 stream
167 dynamic blocks: 3,612,441(5) -> 7,791,488 (0.464)
    lits: 1,871,128(7) -> 1,919,958 (0.975)
    mats: 1,723,913(3) -> 5,871,530 (0.294)
    bits: μ = 173,051.1, σ = 32,053.9, in [42,876..213,484]
    head: μ = 820.3, σ = 210.9, in [249..1,133]
    syms: μ = 16,306.4, σ = 989.4, in [3,597..16,383]
    lens: μ = 7.3, σ = 16.5, in [3..258]
    data: μ = 46,655.6, σ = 28,197.2, in [16,383..236,264]
```

So we can modify firmware

Telnet?

How to draw an owl

1.



2.



1. Draw some circles

2. Draw the rest of the fucking owl

```
$ nc -c 192.168.1.207 59223
**** Welcome to SELPHY TELNET Server ****

login password=alps-alsi
SELPHY_Command>iperf --help
Usage: iperf [-s | -c host] [options]
          iperf --help
Client / Server:
          -p <port>    server port to listen on/connect to
          -u              use UDP
```

Contact

- <https://bsky.app/profile/fuzzyaleks.bsky.social>
- <https://infosec.exchange/@FuzzyAleks>
- <https://github.com/ea/>
- <https://blog.29b.net/>
-

