Week 1a: Understanding the ARM Stack: Inline Assembly and Live Debugging

---

LEGAL DISCLAIMER: The information, tools, and code provided in this repository and course are strictly for educational, research, and defensive purposes only.

You are explicitly prohibited from using any materials contained herein to access, test, modify, or exploit any device, network, or system that you do not own 100% or for which you do not have explicit, documented, and legally binding authorization to interact with.

By using this repository and course, you acknowledge and agree that:

1. Any illegal, unauthorized, or malicious use of this information is solely your responsibility.
2. The author(s) and contributor(s) of this repository and course shall not be held liable for any damages, legal repercussions, criminal charges, or unauthorized actions resulting from the use, misuse, or abuse of the contents herein.
3. You will comply with all applicable local, state, national, and international laws regarding cybersecurity and computer fraud.

IF YOU DO NOT AGREE WITH THESE TERMS, DO NOT USE THIS REPOSITORY AND COURSE.

What You'll Learn This Week

By the end of this week, you will be able to:

- Understand how the RP2350 Cortex-M33 stack grows in SRAM.
- Identify the ARM registers used by this stack experiment.
- Build and flash a Pico 2 ELF through a Debug Probe.
- Connect OpenOCD and GDB to live hardware.
- Step one assembly instruction at a time with `si`.
- Examine the exact stack words written by each multi-register instruction.
- Prove that an ARM register list is ordered by register number, not source-list spelling.

Part 1: Understanding the Basics

What is a Microcontroller?

A microcontroller is a complete small computer on one chip. It contains processor cores, memory controllers, peripherals, and interfaces for hardware such as GPIO, UART, timers, and SPI. The Raspberry Pi Pico 2 uses the **RP2350** microcontroller.

What is the ARM Cortex-M33?

The RP2350 can run Arm Cortex-M33 cores. The program in this folder is built for that Arm target. We will use the Debug Probe, OpenOCD, and GDB to stop a core and inspect its registers and memory while it executes the program.

What is Dynamic Analysis?

Dynamic analysis means observing a program while it runs on real hardware. In this lesson we will:

- Stop the processor at `main`.
- View the instructions produced by the compiler.
- Execute one instruction with `si`.
- Read the stack pointer and the memory it points to.

Part 2: Understanding Processor Registers

What is a Register?

A register is very fast storage inside the CPU. Instructions use registers for values, addresses, temporary results, and control flow.

The ARM Cortex-M33 Registers

Register	Also Called	Purpose
`r0` - `r12`	General purpose	Hold values and addresses while instructions run.
`r13`	SP	Points to the current top of the stack.
`r14`	LR	Holds the return address after a function call.
`r15`	PC	Points to the next instruction to execute.

General-Purpose Registers (`r0` - `r12`)

This lesson uses `r2`, `r3`, `r4`, `r6`, `r9`, and `r10`. The assembly saves their current values to SRAM, then restores them before the loop repeats.

The Stack Pointer (`r13` / SP)

The stack is a region of SRAM used for temporary values, saved registers, return addresses, and local variables. On Cortex-M, the standard stack grows toward lower addresses.

- A `push` lowers `sp` and writes values below the old stack pointer.
- A `pop` reads values at `sp` and raises `sp`.
- Each saved register occupies 4 bytes.

```
Higher addresses
+------------------+
| Old SP location  |
+------------------+
| Saved value      |
+------------------+
| Saved value      |  <- SP after a multi-register push
+------------------+
Lower addresses
```

The Link Register (`r14` / LR)

A `bl` instruction calls a function and stores the return address in `lr`. The compiler-generated prologue for `main` saves `lr` on the stack before calling `stdio_init_all`.

The Program Counter (`r15` / PC)

The Program Counter identifies the next instruction. In GDB, the `=>` marker in `disas main` points to the instruction that will run when you type `si`.

Part 3: Understanding Memory Layout

XIP - Execute In Place

The Pico 2 executes this firmware directly from external flash through XIP. The executable code normally begins at `0x10000000`.

Memory Map Overview

```
+-----------------------------------+
| Flash Memory (XIP)                |
| Starts at: 0x10000000             |
| Contains: program instructions    |
+-----------------------------------+
| SRAM                              |
| Starts at: 0x20000000             |
| Contains: stack, heap, variables  |
+-----------------------------------+
```

Why the Stack Is in SRAM

The stack changes on every function call and return, so it must be writable. When GDB displays `$sp`, it should show an address in the `0x200...` SRAM range. `x/wx $sp` reads the 32-bit value currently at the top of that stack.

Part 3.5: Reviewing Our Stack Code

The file `0x0001a_stack.c` initializes standard I/O, then repeats this assembly block forever:

```
__asm volatile(
    "push {r4, lr}\n"
    "push {r3, r2, r6}\n"
    "stmdb sp!, {r9, r10}\n"
    "ldmia sp!, {r9, r10}\n"
    "pop {r2, r3, r6}\n"
    "pop {r4, lr}\n"
    ::: "memory");
```

Breaking Down the Code

First Push: `push {r4, lr}`

This lowers `sp` by 8 bytes. At the new stack pointer, the saved values are:

```
[sp]     = r4
[sp + 4] = lr
```

Second Push: `push {r3, r2, r6}`

This is deliberately written in a confusing order. The source says `r3` first, but an ARM register list is a set of registers, not an ordered sequence of operations. The assembler encodes the same register mask as `{r2, r3, r6}` and warns that the list is not ascending.

After one `si`, the stack layout proves the actual rule:

```
[sp]     = r2
[sp + 4] = r3
[sp + 8] = r6
[sp + 12] = r4
[sp + 16] = lr
```

The first three words were written by this instruction; `r4` and `lr` remain from the preceding `push {r4, lr}`. The lowest register number in this push is stored at the lowest address. Because the stack grows down, `r6` is closest to the stack pointer value from before this instruction.

High Registers: `stmdb sp!, {r9, r10}`

The 16-bit Thumb `push` encoding cannot encode high registers `r8` through `r12`. `stmdb sp!` is the general full-descending stack instruction used for `r9` and `r10`.

```
[sp]     = r9
[sp + 4] = r10
```

The Restore Instructions

The next instructions restore the same groups in reverse stack-group order. This matters because the last group saved is at the current top of the stack:

```
ldmia sp!, {r9, r10}
pop {r2, r3, r6}
pop {r4, lr}
```

The stack pointer ends at the same value it had before the inline assembly, so the infinite loop does not consume stack space.

Compiling and Flashing to the Pico 2

Step 1: Compile the Code

From the project folder, build the program:

```
& "$env:USERPROFILE\.pico-sdk\ninja\v1.13.2\ninja.exe" -C build
```

The expected artifact is `build\0x0001a_stack.elf`. The assembler reports `register range not in ascending order` for the deliberately unordered source list. That warning is expected for this experiment.

Step 2: Flash and Verify

Use the Debug Probe to flash the ELF and compare target flash with the built image:

```
& "$env:USERPROFILE\.pico-sdk\openocd\0.12.0+dev\openocd.exe" -s "$env:USERPROFILE\.pico-sdk\openocd\0.12.0+dev\scripts" -f interface/cmsis-dap.cfg -f target/rp2350.cfg -c 'adapter speed 5000; targets rp2350.dap.core1; cortex_m reset_config sysresetreq; targets rp2350.dap.core0; program "build/0x0001a_stack.elf" verify reset exit'
```

You should see:

```
** Programming Finished **
** Verify Started **
** Verified OK **
** Resetting Target **
shutdown command invoked
```

`Verified OK` proves that the programmed bytes match the ELF. `shutdown command invoked` is normal because `exit` ends OpenOCD after flashing.

Part 4: Dynamic Analysis with GDB

Prerequisites

Before starting, you need:

1. A Pico 2 with the Debug Probe connected.
2. OpenOCD from the installed Pico SDK.
3. `arm-none-eabi-gdb`.
4. The verified `build\0x0001a_stack.elf` on the Pico 2.

Connecting to Your Pico 2 with OpenOCD

Open a terminal and start the debug server. Leave it running:

```
& "$env:USERPROFILE\.pico-sdk\openocd\0.12.0+dev\openocd.exe" -s "$env:USERPROFILE\.pico-sdk\openocd\0.12.0+dev\scripts" -f interface/cmsis-dap.cfg -f target/rp2350.cfg -c "adapter speed 5000; init"
```

OpenOCD listens for GDB connections on port `3333`.

VM Command

If the VM has `openocd` on its `PATH`, use the same command without the full executable path:

```
openocd -s "$env:USERPROFILE\.pico-sdk\openocd\0.12.0+dev\scripts" -f interface/cmsis-dap.cfg -f target/rp2350.cfg -c "adapter speed 5000; init"
```

The VM needs USB access to the Debug Probe. Nothing else about the build, ELF, or GDB sequence changes.

Connecting to Your Pico 2 with GDB

Open a second terminal in the project folder:

```
arm-none-eabi-gdb build\0x0001a_stack.elf
```

Connect, reset, halt, and stop at `main`:

```
target extended-remote :3333
monitor reset halt
b main
c
disas main
```

You should see this instruction pattern. Your addresses can differ after a rebuild.

```
=> main:          push    {r3, lr}
   main+2:        bl      stdio_init_all
   main+6:        push    {r4, lr}
   main+8:        push    {r2, r3, r6}
   main+10:       stmdb   sp!, {r9, r10}
   main+14:       ldmia.w sp!, {r9, r10}
   main+18:       pop     {r2, r3, r6}
   main+20:       ldmia.w sp!, {r4, lr}
   main+24:       b.n     main+6
```

Notice that GDB displays `{r2, r3, r6}`, not the source spelling `{r3, r2, r6}`. That is the encoded register set in canonical order. Depending on the disassembler, register `r10` may be displayed as its conventional alias, `sl`.

Basic GDB Commands: Your First Steps

Command	Short Form	What It Does
`break main`	`b main`	Set a breakpoint at `main`.
`continue`	`c`	Run until a breakpoint.
`disassemble`	`disas`	Show the current function's assembly.
`info registers`	`i r`	Display CPU registers.
`stepi`	`si`	Execute exactly one instruction.
`nexti`	`ni`	Execute one instruction without entering a call.
`x/wx ADDRESS`		Examine one 32-bit hexadecimal word.
`monitor reset halt`		Ask OpenOCD to reset and halt the target.

Watching the Stack Change

Step 1: Inspect the Stack Before the Compiler Prologue

At the breakpoint, GDB is paused before `push {r3, lr}`. Inspect the current stack pointer and the two words below it:

```
p/x $sp
x/2wx $sp-8
```

Step 2: Execute One Instruction

```
si
```

The arrow moves to `bl stdio_init_all`. Inspect what the compiler prologue placed on the stack:

```
p/x $sp
x/wx $sp
x/wx $sp+4
```

The stack pointer moved down 8 bytes. `[sp]` is the saved `r3`; `[sp+4]` is the saved `lr`.

Step 3: Step Over `stdio_init_all`

Do not step into the library initialization code. Use `ni`:

```
ni
disas main
```

The arrow now points at the first inline instruction: `push {r4, lr}`.

Step 4: Prove the First Inline Push

Read the registers before saving them:

```
p/x $r4
p/x $lr
p/x $sp
```

Execute one instruction and examine the new top of the stack:

```
si
x/wx $sp
x/wx $sp+4
```

The values at `[sp]` and `[sp+4]` match the values shown for `r4` and `lr`. This push reduced `sp` by 8 bytes.

Step 5: Prove Register-List Ordering

Read the three registers before executing the deliberately unordered list:

```
p/x $r2
p/x $r3
p/x $r6
p/x $sp
```

Now execute only that instruction:

```
si
x/wx $sp
x/wx $sp+4
x/wx $sp+8
```

Compare the values from the first three commands with the three words in SRAM:

```
[sp]     matches r2
[sp + 4] matches r3
[sp + 8] matches r6
```

This is the proof. The source ordered the list as `r3`, `r2`, `r6`, but the stack is laid out by ascending register number. A multi-register push is one CPU instruction, so individual transfers inside that instruction cannot be separately stepped.

Step 6: Prove the High-Register Save

Read the values and execute one instruction:

```
p/x $r9
p/x $r10
p/x $sp
si
x/wx $sp
x/wx $sp+4
```

`stmdb sp!, {r9, r10}` moved `sp` down by 8 bytes. The first word equals `r9`; the second equals `r10`.

Step 7: Watch the Restores

Execute and inspect each restore separately:

```
si
p/x $sp
x/3wx $sp

si
p/x $sp
x/2wx $sp

si
p/x $sp
```

The three instructions raise `sp` by 8, 12, and 8 bytes respectively. The final value is the same stack pointer you saw before the first inline push.

Understanding the Stack Diagram

At the deepest point, after `stmdb sp!, {r9, r10}`, the inline assembly has saved 28 bytes. The `old SP` below is the stack pointer after the separate compiler prologue, not the stack pointer at entry to `main`:

```
Before inline assembly:             At maximum inline stack depth:

old SP  <- SP                       old SP
                                     [old SP - 4]  lr
                                     [old SP - 8]  r4
                                     [old SP - 12] r6
                                     [old SP - 16] r3
                                     [old SP - 20] r2
                                     [old SP - 24] r10
                                     [old SP - 28] r9  <- SP
```

The restoration sequence removes the top group first: `r9/r10`, then `r2/r3/r6`, then `r4/lr`.

Part 5: Summary and Review

What We Learned

1. **Registers**: `sp`, `lr`, and `pc` control stack location, returns, and the next instruction.
2. **The stack**: It grows down in SRAM. Pushes decrease `sp`; pops increase it.
3. **Multi-register instructions**: Register-list source order is not the stack-memory order. ARM stores lower register numbers at lower addresses.
4. **OpenOCD and GDB**: OpenOCD connects to the Debug Probe; GDB connects to OpenOCD on port `3333`.
5. **Live evidence**: `si` executes one instruction, and `x/wx $sp` shows the exact 32-bit word that instruction placed at the stack pointer.

The Program Flow

```
+-----------------------------------------------------+
| 1. push {r3, lr}                                    |
|    Compiler saves its prologue registers             |
+-----------------------------------------------------+
| 2. bl stdio_init_all                                 |
|    Initialize standard I/O                           |
+-----------------------------------------------------+
| 3. push {r4, lr}                                    |
|    Save the first inline group                       |
+-----------------------------------------------------+
| 4. push {r3, r2, r6}                                |
|    Source order differs from stack-memory order      |
+-----------------------------------------------------+
| 5. stmdb / ldmia / pop / pop                         |
|    Save and restore all groups                       |
+-----------------------------------------------------+
| 6. b.n main+6                                       |
|    Repeat with the original stack pointer            |
+-----------------------------------------------------+
```

Key Takeaways

1. **The stack grows downward**: a push decreases the numeric value in `sp`.
2. **A register list is not a sequence**: `{r3, r2, r6}` and `{r2, r3, r6}` encode the same register set.
3. **Memory is the proof**: after `si`, inspect `[sp]`, `[sp+4]`, and `[sp+8]` with GDB.
4. **The restore order matters**: always restore the most recently saved group first.
5. **Balanced stack operations are required**: the loop returns `sp` to its starting value on every iteration.

Glossary

Term	Definition
Assembly	Human-readable form of processor instructions.
Breakpoint	A debugger stop point.
Debug Probe	Hardware interface that lets OpenOCD communicate with the target over SWD.
GDB	GNU Debugger, used to inspect and control the running target.
LR	Link Register, holding a function return address.
OpenOCD	Debug server that bridges GDB and the Debug Probe.
PC	Program Counter, pointing to the next instruction.
SP	Stack Pointer, pointing to the top of the stack.
SRAM	Writable memory used for runtime data and the stack.
XIP	Execute In Place, executing program code directly from flash.