## 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

funct7	rs2	rs1	funct3	rd	opcode	R-type
imm[11:	:0]	rs1	funct3	rd	opcode	I-type
imm[11:5]	rs2	rs1	funct3	imm[4:0]	opcode	S-type
imm[12 10:5]	rs2	rs1	funct3	rd	opcode	B-type
	imm[31:12]			rd	opcode	U-type
ir	nm[20 10:1 11 1	9:12]		rd	opcode	J-type

# **Zbb**: "Basic bit-manipulation" Extension

31						25	24		20	19	15 14		12	11	7	6						0	
0	1	0	0	0	0	0	rsi	2		rs1	1	1	1	rd		0	1	1	0	0	1	1	ANDN
0	1	0	0	0	0	0	rsi	2		rs1	1	1	0	rd		0	1	1	0	0	1	1	ORN
0	1	0	0	0	0	0	rsi	2		rs1	1	0	0	rd		0	1	1	0	0	1	1	XNOR
0	1	1	0	0	0	0	0 0 0	0	0	rs1	0	0	1	rd		0	0	1	0	0	1	1	CLZ
0	1	1	0	0	0	0	0 0 0	0	1	rs1	0	0	1	rd		0	0	1	0	0	1	1	CTZ
0	1	1	0	0	0	0	0 0 0	1	0	rs1	0	0	1	rd		0	0	1	0	0	1	1	CP0P
0	0	0	0	1	0	1	rsi	2		rs1	1	1	0	rd		0	1	1	0	0	1	1	MAX
0	0	0	0	1	0	1	rsi	2		rs1	1	1	1	rd		0	1	1	0	0	1	1	MAXU
0	0	0	0	1	0	1	rs	2		rs1	1	0	0	rd		0	1	1	0	0	1	1	MIN
0	0	0	0	1	0	1	rsi	2		rs1	1	0	1	rd		0	1	1	0	0	1	1	MINU
0	1	1	0	0	0	0	0 0 1	0	0	rs1	0	0	1	rd		0	0	1	0	0	1	1	SEXT.B
0	1	1	0	0	0	0	0 0 1	0	1	rs1	0	0	1	rd		0	0	1	0	0	1	1	SEXT.H
0	0	0	0	1	0	0	0 0 0	0	0	rs1	1	0	0	rd		0	1	1	0	0	1	1	ZEXT.H
0	1	1	0	0	0	0	rsi	2		rs1	0	0	1	rd		0	1	1	0	0	1	1	ROL
0	1	1	0	0	0	0	rs	2		rs1	1	0	1	rd		0	1	1	0	0	1	1	ROR
0	1	1	0	0	0	0	shar	nt		rs1	1	0	1	rd		0	0	1	0	0	1	1	RORI
0	0	1	0	1	0	0	0 0 1	1	1	rs1	1	0	1	rd		0	0	1	0	0	1	1	ORC.B
0	1	1	0	1	0	0	1 1 0	0	0	rs1	1	0	1	rd		0	0	1	0	0	1	1	REV8

## 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

funct7	rs2	rs1	funct3	rd	opcode	R-type
imm[11:	0]	rs1	funct3	rd	opcode	I-type
imm[11:5]	rs2	rs1	funct3	imm[4:0]	opcode	S-type
imm[12 10:5]	rs2	rs1	funct3	rd	opcode	B-type
	imm[31:12]			rd	opcode	U-type
in	ım[20 10:1 11 1	9:12]		rd	opcode	J-type

## Zri: "Load/Store indirect with Index" Extension

31						25	24 2	<b>0</b> 19		15	14		12	11	7	6						0	_
0	0	0	0	0	0	0	rs2		rs1		1	1	1	rd		0	0	0	0	0	1	1	LB.R
0	0	0	0	0	0	1	rs2		rs1		1	1	1	rd		0	0	0	0	0	1	1	LH.R
0	0	0	0	0	1	0	rs2		rs1		1	1	1	rd		0	0	0	0	0	1	1	LW.R
1	0	0	0	0	0	0	rs2		rs1		1	1	1	rd		0	0	0	0	0	1	1	LBU.R
1	0	0	0	0	0	1	rs2		rs1		1	1	1	rd		0	0	0	0	0	1	1	LHU.R
0	0	0	0	0	0	0	rs3		rs1		1	1	1	rs2		0	1	0	0	0	1	1	SB.R
0	0	0	0	0	0	1	rs3		rs1		1	1	1	rs2		0	1	0	0	0	1	1	SH.R
0	0	0	0	0	1	0	rs3		rs1		1	1	1	rs2		0	1	0	0	0	1	1	SW.R

1b rd, rs2(rs1)

lb rd, rs2(rs1)
lh rd, rs2(rs1)
lw rd, rs2(rs1)
lbu rd, rs2(rs1)
lhu rd, rs2(rs1)
sb rs2, rs3(rs1)
sh rs2, rs3(rs1)
sw rs2, rs3(rs1)

## 31 30 29 28 27 26 25 24 23 22 21 20 19 18 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

funct7	rs2	rs1	funct3	rd	opcode	R-type
imm[11:	0]	rs1	funct3	rd	opcode	I-type
imm[11:5]	rs2	rs1	funct3	imm[4:0]	opcode	S-type
imm[12 10:5]	rs2	rs1	funct3	rd	opcode	B-type
	imm[31:12]			rd	opcode	U-type
im	m[20 10:1 11 1	9:12]		rd	opcode	J-type

# Zor: "Objective RISC" Extension

# <u>Unprivileged:</u>

<u> </u>	<u> </u>																									
31						25	24				20	19		15	14		12	11 7	6						0	_
0	0	0	0	0	0	0			rs2				rs1		0	0	0	rs3	0	0	0	1	0	1	1	SP.R
0	0	0	0	0	0	1			rs2				rs1		0	0	0	rd	0	0	0	1	0	1	1	LP.R
0	0	0	0	0	1	0	i	inde	ex[4	4:0	]		frame	)	0	0	0	rs1	0	0	0	1	0	1	1	SV
0	0	0	0	0	1	1	i	inde	ex[4	4:0	]		frame	ì	0	0	0	rd	0	0	0	1	0	1	1	RST
0	0	0	0	1	0	0		7	zer	О			rs1		0	0	0	rd	0	0	0	1	0	1	1	QDTB
0	0	0	0	1	0	1		Z	zer	)			rs1		0	0	0	rd	0	0	0	1	0	1	1	QDTH
0	0	0	0	1	1	0		7	zer	)			rs1		0	0	0	rd	0	0	0	1	0	1	1	QDTW
0	0	0	0	1	1	1		2	zer	Э			rs1		0	0	0	rd	0	0	0	1	0	1	1	QDTD
0	0	0	1	0	0	0		7	zer	)			rs1		0	0	0	rd	0	0	0	1	0	1	1	QPI
0	0	0	1	0	0	1		2	zer	О			zero		0	0	0	rd	0	0	0	1	0	1	1	GCP
0	0	0	1	1	0	0		2	zer	О			frame	h.	0	0	0	frame	0	0	0	1	0	1	1	POP
0	0	1	0	0	0	1		2	zen	О			zero		0	0	0	zero	0	0	0	1	0	1	1	RTLIB
0	0	1	0	0	1	0		- 2	zer	0			zero		0	0	0	zero	0	0	0	1	0	1	1	CPFC
0	0	1	0	0	1	1		2	zer	0			zero		0	0	0	zero	0	0	0	1	0	1	1	CHECK
		imm	[11	_	_				rs2				rs1		0	0	1	imm[4:0]	0	0	0	1	0	1	1	SP
				ir	nm [ :	11:	0]						rs1		0	1	0	rd	0	0	0	1	0	1	1	LP
				ir	nm [	11:	0]						rs1		0	1	1	ra	0	0	0	1	0	1	1	JLIB
0	0	0	0	0	0	0			rs2				rs1		1	0	0	rd	0	0	0	1	0	1	1	ALC
				р	i[1	1:6	9]						rs1		1	0	1	rd	0	0	0	1	0	1	1	ALCI.P
				d	t[1	1:6	9]						rs1		1	1	0	rd	0	0	0	1	0	1	1	ALCI.D
		dt	[6:	0]			0	0	0	0	0		rd		1	1	1	pi[4:0]	0	0	0	1	0	1	1	ALCI
		dt	[6:	0]			0	0	0	1	0		frame		1	1	1	pi[4:0]	0	0	0	1	0	1	1	PUSHG
		dt	[6:	0]			0	0	0	1	1		frame	h.	1	1	1	pi[4:0]	0	0	0	1	0	1	1	PUSH

## Machine Mode:

31					26	25	24				20	19				15	14		12	11	7	6						0	_
1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	rd		1	1	1	0	0	1	1	ALCB
1	1	1	1	1	1	1			rs2					rs1	L		0	0	0	rd		1	1	1	0	0	1	1	CIOP
1	1	1	1	1	1	0	1	0	0	0	0			rs1	L		0	0	0	rd		1	1	1	0	0	1	1	CCP
1	1	1	1	1	1	0	1	0	0	0	1			rs1	L		0	0	0	rd		1	1	1	0	0	1	1	RPR
1	1	1	1	1	1	0	1	0	1	0	0			rs1	L		0	0	0	rd		1	1	1	0	0	1	1	QPIR
1	1	1	1	1	1	0	1	0	1	0	1			rs1	L		0	0	0	rd		1	1	1	0	0	1	1	QDTR
1	1	1	1	1	1	0	1	0	1	1	0			rs1	L		0	0	0	rd		1	1	1	0	0	1	1	QPTR
1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0	rd		1	1	1	0	0	1	1	SEAL
1	1	1	1	1	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	rd		1	1	1	0	0	1	1	UNSL

## Misc:

11130.			
reg	alias	reg	alias
х0	zero	x16	a6
x1	ra <del>rix</del>	x17	a7
x2	frame	x18	s2
х3	<del>rcd/</del> root/core	x19	s3
x4	ctxt	x20	s4
x5	t0	x21	s5
х6	t1	x22	s6
x7	t2	x23	s7
x8	s0	x24	s8
x9	s1	x25	s9
x10	a0	x26	s10/bm
x11	a1	x27	cnst
x12	a2	x28	t3
x13	a3	x29	t4
x14	a4	x30	t5
x15	a5	x31	t6

pseudo-instruction	implemented as
lcp rd, imm(rs1)	lp rd, imm(rs1)
	sp x0, imm(rs1)
lcp.r rd, imm(rs1)	lp.r rd, rs2(rs1)
	sp.r x0, rs2(rs1)
scp rs2, imm(rs1)	sp rs2, imm(rs1)
	addi rs2, x0,0
scp.r rs2, rs3(rs1)	sp.r rs2, rs3(rs1)
	addi rs2, x0,0
pusht pi,dt	alci frame, pi,dt

R R R R R R R R R R R I I S S S

R R R R R R R R

# Implementation:

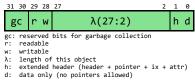
Instruction	rdst	rdat	rptr	raux	imm
sb/h/w	zero	ra.rix	rs1	rs2	imm
lb/bu/h/hu/w	rd		rs1	ra	imm
sp	zero	ra.rix	rs1	rs2	imm
lp	rd		rs1	ra	imm
sb/h/w.r	zero	rs3	rs1 (≠ frame)	rs2	
lb/bu/h/hu/w.r	rd	rs2	rs1 (≠ frame)		
sp.r	zero	rs3	rs1 (≠ frame)	rs2	
lp.r	rd	rs2	rs1 (≠ frame)		
SV	zero	ra.rix	frame	rs1	index
rst	rd	ra.rix	frame	bm	index
qdtx					
qpi					
gcp					
рор	frame	ra.rix	frame		
jlib	ra	frame	rs1	ra	imm
jal	rd	frame		ra	imm
jr	rd	frame	rs1	ra	imm
rtlib	ra	ra.rix	ra	frame	
alc	rd (≠ frame)	rs1	alc_params	rs2	
alci.p	rd (≠ frame)	rs1	alc_params		pi
alci.d	rd (≠ frame)	rs1	alc_params		dt
alci	rd	ra.rix	alc_params	frame	pi & dt
pushg	rd	ra.rix	alc_params	frame	pi & dt
push	rd	ra.rix	alc_params	frame	pi & dt
alcb					
ciop	rd	rs1		rs2	
rpr					
qpir					
qdtr					
qptr					
seal					
unsl					

	31	30	29	3	2	1	0
ra.rix	lib entry		rix(30:1)				color
frame			frame(31:3)		1	0	color
pi	uini		pi(30:2)			pnwber/gc	gc
dt	rc	ri	dt(29:0)				

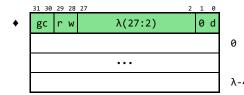
instruction	condition	action
jlib	ra.rix(color) != frame(color)	set ra.rix(lib entry), toggle rix(color)
	target ptr != ra.rcd	
jal ra, or jr ra,	ra.rix(color) != frame(color)	clear ra.rix(lib entry), toggle rix(color)
pushx	ra.rix(color) = frame(color)	toggle frame(color)
рор	ra.rix(color) != frame(color)	toggle frame(color)
jr, 0(ra)	ra.rix(color) = frame(color)	toggle ra.rix(color)
		if ra.rix(lib entry) = 1 do cross code-object return
		else stay in this code-object

## **OBJECTS**

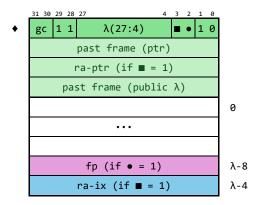
#### Generic Header



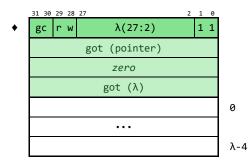
## Ordinary



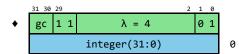
#### Frame



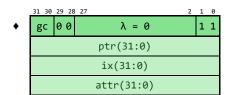
## Executable



## Immediate (Primitive)

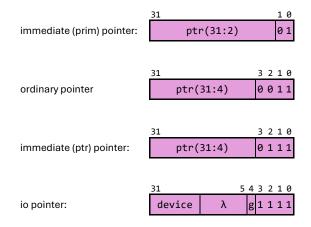


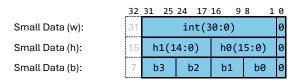
## Immediate (Pointer)



# **POINTERS & DATA**

(in memory)





Allocate immediate primitive if:

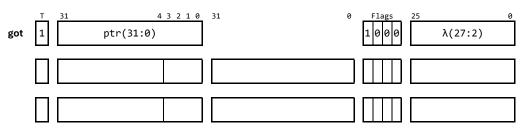
- sw and rs(30) ≠ rs(31)
- sh at h1 and rs(14) ≠ rs(15)
- sb at b3 and (rs(7) = 1 or rs < 0)

# **REGISTER FILE & PIPELINE**

# Architectural Registers (x0-x31, alc-params, nano-scratch):

data	т Ø	value(31:		alc_addr	0	Flags 0 0 0	25	zero 2	1 0
ordinary pointer	1	ptr(31:4)	0000	index(31:0)		rwxd		λ(27:2)	
frame-type	1	ptr(31:4)	1 1 1 c	index(31:0)		1 1 0 0		λ(27:4)	•
ra	1	ptr(31:4)	1 1 0 c	index(31:0)		r w 1 1		λ(27:2)	
io pointer	1	ptr(31:4)	1000	index(31:0)		r w 0 1		λ(27:2)	
nano-scratch (nscr)	0	value(31:	0)						
alc-params (ap)	1	alc. addr	· .	alc. lim.					

# **Microarchitectural Registers:**



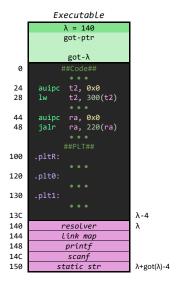
# **FRAME OPERATIONS**

We support 4 different ways to use the stack:

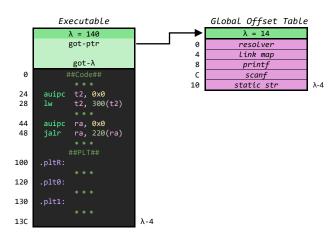
Standard RISC-V ABI access to the past stack frame allowed via sp and fp	Standard RISC-V ABI (-fno-omit-frame-pointer) access to the past stack frame allowed only via fp	use fp for parameters access to the past stack frame only with explicit ptr (not sp)	split frame in private and public μ-frames access to the past stack frame only with explicit ptr (not sp)
routine0: addi sp, sp,-48 sw ra, 44(sp) sw s1, 40(sp)  li t0, 8 sw t0, 16(sp)  li a7, 3 sw a7, 4(sp) addi a7, sp,24 sw a7, 0(sp) jal ra, routine1  routine1: addi sp, sp,-16 sw ra, 12(sp)	routine0:    addi   sp, sp,-48    sw   ra, 44(sp)    sw   s0, 40(sp)    addi   s0, sp,48    sw   s1, 36(sp)  li   t0, 8    sw   t0, -32(s0)  li   a7, 3    sw   a7, 4(sp)    addi   a7, s6,-28    sw   a7, 0(sp)    jal   ra, routine1  routine1:	routine0:    addi   sp, sp,-32    sw   ra, 28(sp)    sw   s0, 24(sp)    sw   s1, 20(sp)  li   t0, 8    sw   t0, 0(sp)  alci   s0, 8    li   a7, 3    sw   a7, 4(s0)    addi   a7, sp,4    sw   a7, 0(s0)    jal   ra, routine1  routine1:	routine0:    addi   sp, sp,-16    sw   ra, 12(sp)    sw   s0, 8(sp)    alci   s0, 20    sw   s1, 4(sp)  li   t0, 8    sw   t0, 0(sp)  li   a7, 3    sw   a7, 4(s0)    addi   a7, s0,8    sw   a7, 0(s0)    jal   ra, routine1  routine1:
<pre>lw a1, 16(sp) lw a0, 0(a1) lw t0, 20(sp) add a1, a1,t0 lw a1, 0(a1) jal ra, routine2 routine2: add a0, a0,a1 ret</pre>	addi sp, sp,-16 sw ra, 12(sp) sw s0, 8(sp) addi s0, sp,16  lw a1, 0(s0) lw a0, 0(a1) lw t0, 4(s0) add a1, a1,t0 lw a1, 0(a1) jal ra, routine2  routine2: add a0, a0,a1 ret	addi sp, sp,-16 sw ra, 12(sp) sw s0, 8(sp) addi s0, sp,16  lw a1, 0(s0) lw a0, 0(a1) lw t0, 4(s0) add a1, a1,t0 lw a1, 0(a1) jal ra, routine2  routine2: add a0, a0,a1 ret	addi sp, sp,-16 sw ra, 12(sp)  lw a1, 0(s0) lw a0, 0(a1) lw t0, 4(s0) add a1, a1, t0 lw a1, 0(a1) jal ra, routine2  routine2: add a0, a0,a1 ret
sp• 8b	\$b\$  3  4  8  12  int a  16  int b[0]  int b[1]  int b[2]  28  int b[3]  31  32  \$1  \$0  44  \$0	sp> int a	spr int a
sp• 0 4 8 12	spr 0 4 4 s0 8 ra 12 s0r	spr 0 4 50 8 ra 12	sp* 0 4 8 12

# **CODE SEGMENTATION**

Virtual Structure



## Actual Structure



## **User Mode Instructions**

Instruction		rd	rs1	rs2	cr	imm	Decoder Decision
lui		rd			-	imm	
aui	рс	rd			-	imm	
jal		rd		sp	•	imm	
jal	r	rd	rs1		-	imm	
Α	jalr	rd	rs1	sp	•	imm	always
Α	lgt	got	rs1		-		always (instead of nop)
bcc			rs1	rs2	-	imm	
1b/	bu/h/hu/w	rd	rs1		-		
Α	lb/bu/h/hu/w	rd	rs1		•		if rs1 = fp or rs1 = sp
В	lb/bu/h/hu/w	rd	rs1	got	•		otherwise
sb/			rs1	rs2	-	imm	
Α	loadmux	scr	rs1	rs2	•	imm	if sb and imm(0) = 1
Α	sb_m/h_m	ар	rs1	scr	•	imm	or sh and imm(1) = 1
В	sb/h/w	ар	rs1	rs2	•	imm	otherwise
add	i	rd	rs1		-	imm	
Α	push	sp	sp		•	imm	if rd = sp and rs1 = sp and imm > 0
В	рор	sp	sp		-	imm	if rd = sp and rs1 = sp and imm < 0
С	setpublic	zero	sp		•	imm	if rd = zero and rs1 = sp and imm ≥ 0
D	addi	rd	rs1		-	imm	otherwise
ari	thi	rd	rs1		-	imm	
ari	th	rd	rs1	rs2	-		
alc		rd	rs1	ар	-		
alc		rd		ар	-	imm	
alc		rd	rs1	ар	-		
alc	i.d	rd		ар	-	imm	
qsz		rd	rs1		-		

# **Supervisor Mode Instructions:**

Instruction	rd	rs1	rs2	cr	imm	Notes
sb/h/w.r		rs1	rs2	-	imm	"store raw", allows stores at any point in memory. Uses rs1 as base-ptr
lb/bu/h/hu/w.r	rd	rs1		-		"load raw", same as store raw
dtp	rd	rs1		-		"data to pointer", creates a pointer from data
ptd	rd	rs1		-		"pointer to data", extracts base address of pointer as data
itd	rd	rs1		-		"index to data", extracts index of pointer as data

## **DOKUMENTATION: ELF-FILES**

"Executable and Linkable Format"-Files bestehen mindestens aus einem Header, einer "Program Header Table" und einer "Section Header Table". Im Header werden Informationen über das ELF-File selbst gespeichert, wie z.B. die Prozessorarchitektur, für welche das Programm kompiliert wurde und die Positionen der PHT und der SHT in Relation zum File-Anfang. In einem Program Header werden Informationen gespeichert, die dem Betriebssystem angeben, wie viele und welche Arten von virtuellen Seiten für dieses Programm benötigt werden. In einem Section Header wird angegeben, in welche Einzelteile das Programm zerlegt wurde und ob noch mehr Informationen über das Programm im ELF-File zu finden sind (z.B. für relocatable Programme).

## **Daten**

Statische Daten werden von einem Compiler über Assemblerdirektiven immer so in die .data bzw. .rodata Sektionen abgelegt, sodass sie in der Symboltabelle des ELF-Files immer als Objekt mit seiner Größe eindeutig erkennbar sind.

```
static char stringA[] = "hello world!";
                                                               static const char stringB[] = "hello world!";
         stringA, @object
.asciz "hello world!"
                                                                         stringB, @object
.asciz "hello world!"
stringA: .asciz
                                                               stringB: .asciz
          stringA, .-stringA
                                                                          stringB, .-stringB
Section Headers:
                                         Offset
  [Nr] Name
                                                                EntSize
                                                                            Flags Link Info Align
                              Address
                                                     Size
  [ 5] .data
                   PROGBITS 00002010 000003b4 0000000d
                                                                00000000
                                                                             WA
                                                                                     0
       .rodata
                   PROGBITS 00002020 000003c4
                                                     00000000
                                                                00000000
                                                                              Α
                                                                                     0
                                                                                                4
//Symbol Table im erzeugten ELF-File
Symbol table '.symtab' contains 60 entries:
                                    Bind
   Num: Value
                     Size Type
                                                      Ndx Name
    49: 00000000
                       13 OBJECT LOCAL
                                          DEFAULT
                                                       5 stringA
    50: 00000000
                       13 OBJECT LOCAL
                                          DEFAULT
                                                       6 stringB
```

Ein Zugriff auf solche statischen Daten kann in executables und muss in relocatables über die Global Offset Table (GOT) stattfinden. Angenommen ein Programm läge an der physikalischen Adresse 0x0 und seine zugehörige GOT an der Adresse 0x1000 und am Offset 8 der GOT stünde die Adresse für das Symbol stringA, dann würde mit folgenden Assembly befehlen auf diesen Eintrag zugegriffen werden.

```
auipc t2, 0x1 # R_RISCV_GOT_HI20 (symbol), R_RISCV_RELAX
lw t2, 8(t2) # R_RISCV_PCREL_LO12_I (auipc), R_RISCV_RELAX
```

In einer executable können die Immediates für diese Befehlssequenz direkt befüllt werden, da der Abstand des Programms zur GOT schon beim Kompilieren des Programms bekannt ist. Bei einem relocatable Programm belässt der Compiler diese Immediates mit 0 und markiert die Befehle in der "Relocation Section" als unaufgelöst. Sowohl die GOT als auch die .data oder .rodata Sektionen können vom Betriebssystem beim Laden des Programms an beliebige Stellen im Speicher platziert werden. Sind alle Sektionen platziert, kann der Dynamische Linker anhand der Tags der Einträge in der Relocation Section herausfinden, wie er die Immediates für die aufzulösenden Symbole zu berechnen hat. R\_RISCV\_GOT\_HI20 z.B. bedeutet, dass für diese Instruktion die obersten 20 Bits der Differenz aus Position der Instruktion und Position der GOT benötigt. Die Relax Tags sollen anzeigen, dass es je nach Positionierung möglich sein könnte, eine der beiden Instruktionen zu sparen falls z.B. Instruktion und GOT nah genug beieinander liegen.

## Code

Bla bla Procedure Linkage Table