

Threads & GC in Clozure CL

R. Matthew Emerson rme@clozure.com
Clozure Associates

Threads

- Threads in CCL are scheduled by the operating system (so they can run concurrently on multiple processors).
- A thread can get preempted at any instruction boundary; this implies that a GC may happen at any instruction boundary.
- WITHOUT-INTERRUPTS doesn't affect scheduling.

Closure CL GC

- The GC is precise: it always knows whether a root (register or stack location) contains a lisp object or just raw bits.
- Strict register (and stack) usage conventions enable this.
- Registers and stacks must be in a GC-consistent state at every instruction boundary (except for few special cases).

Register Conventions

PowerPC

rzero	fn
sp	temp3
target-1	temp2
imm0	temp1
imm1	temp0
imm2	arg_x
imm3	arg_y
imm4	arg_z
imm5	save7
allocptr	save6
allocbase	save5
nargs	save4
tsp	save3
target-2	save2
loc-pc	save1
vsp	save0

- Some registers always contain “immediates”
- Others always contain “nodes”
- Immediates must never end up in node registers (and vice versa), not even for a single instruction.

Register usage conventions work best when there are registers to work with...

PowerPC

rzero	fn
sp	temp3
target-1	temp2
imm0	temp1
imm1	temp0
imm2	arg_x
imm3	arg_y
imm4	arg_z
imm5	save7
allocptr	save6
allocbase	save5
nargs	save4
tsp	save3
target-2	save2
loc-pc	save1
vsp	save0

x86-64

imm0
imm2
imm1
temp0
rsp
rbp
arg_z
arg_y
arg_x
temp1
temp2
save3
save2
fn
save1
save0

x86-32

imm0
temp0
temp1
arg_z
esp
ebp
arg_y
fn

x86-32 hackery

- A bit mask in thread-private memory indicates whether each register is a node or an immediate.
- We also use the x86 direction flag: if set, EDX is an immediate, otherwise it's a node.

x86-32

eax	imm0
ecx	temp0
edx	temp1
ebx	arg_z
	esp
	ebp
esi	arg_y
edi	fn

Sample code...

```
(defx8632lapfunction %atomic-incf-ptr ((ptr arg_z))
  (mark-as-imm temp0)
  (mark-as-imm temp1)
  (let ((imm1 temp0)
        (imm2 temp1))
    (macptr-ptr ptr imm2)
    @again
    (movl (@ (% imm2)) (% eax))
    (lea (@ 1 (% eax)) (% imm1))
    (lock)
    (cmpxchgl (% imm1) (@ (% imm2))))
    (jne @again)
    (box-fixnum imm1 arg_z))
  (mark-as-node temp0)
  (mark-as-node temp1)
  (single-value-return))
```

So, who cares?

- People hacking the runtime and the compiler back end
- LAP programmers
- The Lisp programmer is not affected by any of these issues.