

变量分配在stack上还是heap上，对性能是会有影响的。

Go初学者容易有一种误解：

- 认为函数参数和返回值如果使用变量的值会对整个变量做拷贝，速度慢
- 认为函数参数和返回值如果使用指针类型，只需要拷贝内存地址，速度更快

通过逃逸分析，**编译器会尽可能把能分配在栈上的对象分配在栈上，避免堆内存频繁GC垃圾回收带来的系统开销，影响程序性能。**

**一般而言，遇到以下情况会发生逃逸行为，Go编译器会将变量存储在heap上**

- 函数内局部变量在函数外部被引用
- 接口(interface)类型的变量
- size未知或者动态变化的变量，如slice, map, channel, []byte等
- size过大的局部变量，因为stack内存空间比较小。

在 Go 项目的开发过程中，我们常会遇到一些令人困惑的性能问题：

- 明明是小对象，GC 却频繁触发
- 某个调用频率极高的函数突然引发内存暴涨
- 程序未运行多久就耗尽了所有可用内存

这些问题，往往可以从 Go 语言的“逃逸分析”中找到答案。

- 栈分配：速度快，成本低，随函数返回自动回收
- 堆分配：需要垃圾回收，有额外开销，但生命周期不受限于函数调用

内存逃逸可能引发的问题在于：当出现大量“由栈到堆”的内存逃逸情况时，会加大垃圾回收（GC）的压力，进而导致性能下降。

## 案例1

我们看下面的代码：其中结构体foo的可以参考 **this example[4]**。

```
1 func getFooValue() foo {
2     var result foo
3     // Do something
4     return result
5 }
6
```

变量result定义的时候会在这个goroutine的stack上分配result的内存空间。

当函数返回时，getFooValue的调用方如果有接收返回值，那result的值会被拷贝给对应的接收变量。

stack上变量result的内存空间会被释放(标记为不可用, 不能再被访问, 除非这块空间再次被分配给其它变量)。

**注意:** 本案例的结构体foo占用的内存空间比较小, 约0.3KB, goroutine的stack空间足够存储, 如果foo占用的空间过大, 在stack里存储不了, 就会分配内存到heap上。

## 案例2 指针逃逸

我们看下面的代码:

```
1 func getFooPointer() *foo {
2     var result foo
3     // Do something
4     return &result
5 }
6
```

函数getFooPointer因为返回的是一个指针, 如果变量result分配在stack上, 那函数返回后, result的内存空间会被释放, 就会导致接受函数返回值的变量无法访问原本result的内存空间, 成为一个悬浮指针(dangling pointer)。

所以这种情况会发生内存逃逸, result会分配在heap上, 而不是stack上。

## 案例3 闭包

我们看下面的代码:

```
1 func main() {
2     p := &foo{}
3     f(p)
4 }
5
```

指针变量p是函数f的实参, 因为我们是在main所在的goroutine里调用函数f, 并没有跨goroutine, 所以指针变量p分配在stack上就可以, 不需要分配在heap上。

## 内存逃逸分析工具

因为内存逃逸分析是编译器在编译期就完成的, 可以使用以编译下命令来做内存逃逸分析:

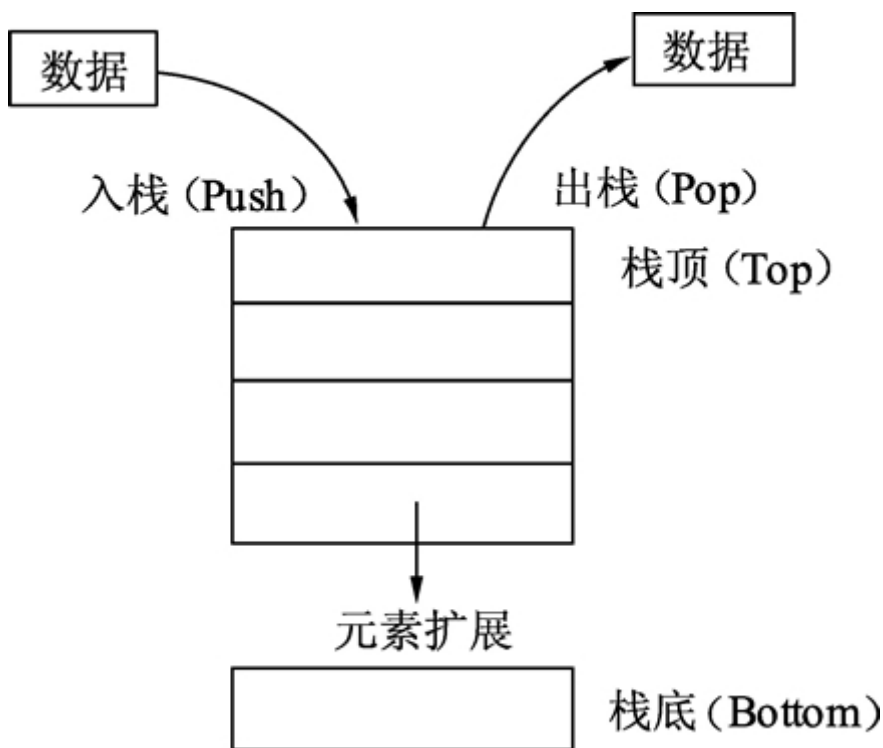
- go build -gcflags="-m", 可以展示逃逸分析、内联优化等各种优化结果。

- `go build -gcflags="-m -l"`，`-l`会禁用内联优化，这样可以过滤掉内联优化的结果展示，让我们可以关注逃逸分析的结果。
- `go build -gcflags="-m -m"`，多一个`-m`会展示更详细的分析结果。

在讨论变量生命周期之前，先来了解下计算机组成里两个非常重要的概念：堆和栈。变量的生命周期

## 什么是栈（stack）

1) 概念栈只允许从线性表的同一端放入和取出数据，按照后进先出（LIFO，Last In First Out）的顺序，如下图所示。



图：栈的操作及扩展

往栈中放入元素的过程叫做入栈。入栈会增加栈的元素数量，最后放入的元素总是位于栈的顶部，最先放入的元素总是位于栈的底部。

从栈中取出元素时，只能从栈顶部取出。取出元素后，栈的元素数量会变少。最先放入的元素总是最后被取出，最后放入的元素总是最先被取出。不允许从栈底获取数据，也不允许对栈成员（除了栈顶部的成员）进行任何查看和修改操作。

2) 变量和栈有什么关系栈可用于内存分配，栈的分配和回收速度非常快。

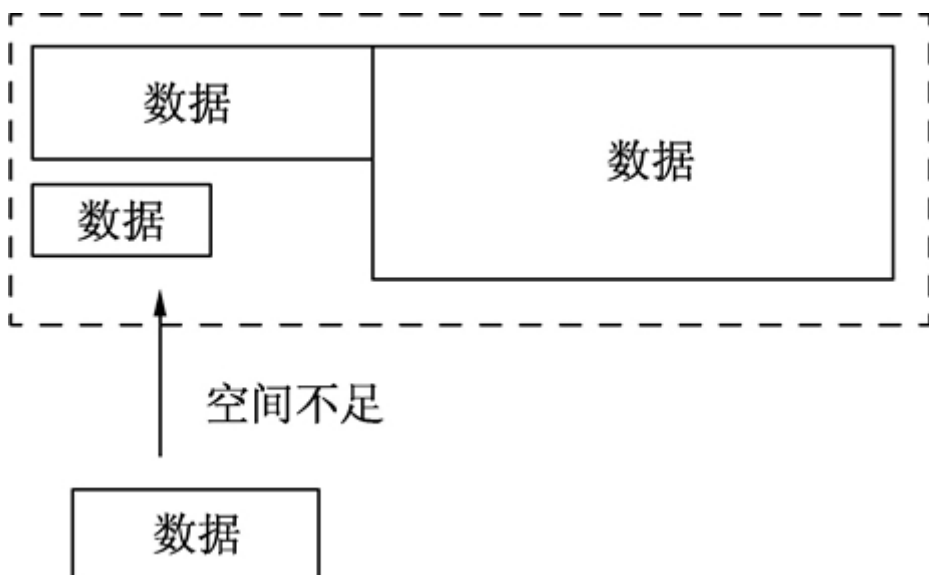
下面的代码展示了栈在内存分配上的作用：

```
1 func calc(a, b int) int {  
2     var c int // 声明整型变量 c, 运行时, c 会分配一段内存用以存储 c 的数值。  
3     c = a * b  
4  
5     var x int // 声明整型变量 x, x 也会被分配一段内存。  
6     x = c * 10  
7  
8     return x  
9 }
```

上面的代码在没有任何优化的情况下，会进行变量 `c` 和 `x` 的分配过程。Go语言默认情况下会将 `c` 和 `x` 分配在栈上，这两个变量在 `calc()` 函数退出时就不再使用，函数结束时，保存 `c` 和 `x` 的栈内存再出栈释放内存，整个分配内存的过程通过栈的分配和回收都会非常迅速。

## 什么是堆 heap

堆在内存分配中类似于往一个房间里摆放各种家具，家具的尺寸有大有小，分配内存时，需要找一块足够装下家具的空间再摆放家具。经过反复摆放和腾空家具后，房间里的空间会变得乱七八糟，此时再往这个空间里摆放家具会发现虽然有足够的空间，但各个空间分布在不同的区域，没有一段连续的空间来摆放家具。此时，内存分配器就需要对这些空间进行调整优化，如下图所示。



图：堆的分配及空间

堆分配内存和栈分配内存相比，堆适合不可预知大小的内存分配。但是为此付出的代价是分配速度较慢，而且会形成内存碎片。

## 变量逃逸 (Escape Analysis) ——自动决定变量分配方式，提高运行效率

堆和栈各有优缺点，该怎么在编程中处理这个问题呢？在 C/C++ 语言中，需要开发者自己学习如何进行内存分配，选用怎样的内存分配方式来适应不同的算法需求。比如，函数局部变量尽量使用栈，全局变量、结构体成员使用堆分配等。程序员不得不花费很长的时间在不同的项目中学习、记忆这些概念并加以实践和使用。

Go语言将这个过程整合到了编译器中，命名为“变量逃逸分析”。通过编译器分析代码的特征和代码的生命周期（根据变量的使用范围和传递方式），判断该变量应该分配在栈上还是堆上。这一过程称为“逃逸分析”。

### 1) 逃逸分析

通过下面的代码来展现Go语言如何使用命令行来分析变量逃逸，代码如下：

```

1 package main
2
3 import "fmt"
4
5 // dummy() 函数拥有一个参数，返回一个整型值，用来测试函数参数和返回值分析情况。
6 func dummy(b int) int {
7
8     // 声明一个变量c并赋值
9     var c int //声明变量 c，用于演示函数临时变量通过函数返回值返回后的情况。
10    c = b
11
12    return c
13 }
14
15 // 空函数，什么也不做
16 func void() {
17 }
18
19 func main() {
20
21    // 声明a变量并打印
22    var a int
23
24    // 调用void()函数 没有返回值，测试 void() 调用后的分析情况
25    void()
26    // 打印 a 和 dummy(0) 的返回值，测试函数返回值没有变量接收时的分析情况。
27    fmt.Println(a, dummy(0))
28 }

```

接着使用如下命令行运行上面的代码：

**go run -gcflags "-m -l" main.go**

使用 go run 运行程序时，-gcflags 参数是编译参数。其中 -m 表示进行内存分配分析，-l 表示避免程序内联，也就是避免进行程序优化。

运行结果如下：

**# command-line-arguments**

**./main.go:29:13: a escapes to heap**

./main.go:29:22: dummy(0) escapes to heap

./main.go:29:13: main ... argument does not escape

0 0程序运行结果分析如下：

- 第 2 行告知“代码的第 29 行的变量 **a 逃逸到堆**”。
- 第 3 行告知“dummy(0) 调用逃逸到堆”。由于 dummy() 函数会返回一个整型值，这个值被 fmt.Println 使用后还是会在 main() 函数中继续存在。
- 第 4 行，这句提示是默认的，可以忽略。

上面例子中变量 c 是整型，其值通过 dummy() 的返回值“逃出”了 dummy() 函数。**变量 c 的值被复制并作为 dummy() 函数的返回值返回**，即使变量 c 在 dummy() 函数中分配的内存被释放，也不会影响 main() 中使用 dummy() 返回的值。**变量 c 使用栈分配不会影响结果。**

## 2) 取地址发生逃逸

下面的例子使用结构体做数据，来了解结构体在堆上的分配情况，代码如下：

```
1 package main
2
3 import "fmt"
4
5 // 声明空结构体测试结构体逃逸情况
6 type Data struct {
7 }
8 // 将 dummy() 函数的返回值修改为 *Data 指针类型。
9 func dummy() *Data {
10     // 实例化c为Data类型
11     var c Data
12
13     //返回函数局部变量地址
14     return &c
15 }
16
17 func main() {
18     fmt.Println(dummy())
19 }
```

执行逃逸分析：

go run -gcflags "-m -l" main.go

# command-line-arguments

./main.go:15:9: &c escapes to heap

./main.go:12:6: moved to heap: c

./main.go:20:19: dummy() escapes to heap

./main.go:20:13: main ... argument does not escape

&{}注意第 4 行出现了新的提示：**将 c 移到堆中**。这句话表示，Go 编译器已经确认如果将变量 c 分配在栈上是无法保证程序最终结果的，如果这样做，dummy() 函数的返回值将是一个不可预知的内存地址，这种情况一般是 C/C++ 语言中容易犯错的地方，引用了一个函数局部变量的地址。

Go语言最终选择将 c 的 Data 结构分配在堆上。然后由垃圾回收器去回收 c 的内存。

3) 原则在使用Go语言进行编程时，Go语言的设计者不希望开发者将精力放在内存应该分配在栈还是堆的问题上，编译器会自动帮助开发者完成这个纠结的选择，

编译器觉得变量应该分配在堆和栈上的原则是：

- 变量是否被取地址；
- 变量是否发生逃逸。

## 指针逃逸

指针逃逸是指函数中创建了一个对象，返回了这个对象的指针。这种情况下，函数虽然退出了，但是因为指针的存在，对象的内存不能随着函数结束而回收，因此只能分配在堆上。

```
1 package main
2
3 type User struct {
4     Name string
5 }
6
7 func GetUser() *User {
8     u := User{Name: "Alice"}
9     return &u // u 逃逸到堆上
10 }
11
12 func main() {
13     u := GetUser()
14     println(u.Name)
15 }
```



## 分析结果

```
1 > go build -gcflags="-m -l" escape1.go
2 # command-line-arguments
3 ./escape1.go:8:2: moved to heap: u
```

## 建议调整为

```
1 func Foo() User {
2     u := User{Name: "Alice"}
3     return u
4 }
```

## interface{} 动态类型逃逸

Go 中将具体类型转为

interface{} 时，如果该值无法在栈上安全传递，也会被转移到堆上。

```
1 package main
2
3 func getMsg() interface{} {
4     return 'a'
5 }
6
7 func main() {
8     m := getMsg()
9     println(m)
10 }
11
```

## 分析结果

```
1 ./point.go:4:9: 'a' escapes to heap
```

## 优化后代码

```
1 func getMsg() msg {  
2     return msg{ }  
3 }  
4  
5 func main() {  
6     m := getMsg()  
7     println(m)  
8 }
```

## 栈空间不足，创建大对象

由于栈空间有限，当创建大对象时，可能会直接分配到堆。

```
1 func generateStack() {  
2     // 创建一个容量为8192的整数切片，其占用内存 <= 64KB  
3     nums := make([]int, 8192)  
4     for i := 0; i < 8192; i++ {  
5         nums[i] = rand.Int()  
6     }  
7 }  
8  
9 func generateHeap() {  
10    // 创建一个容量为8193的整数切片，其占用内存 > 64KB  
11    nums := make([]int, 8193)  
12    for i := 0; i < 8193; i++ {  
13        nums[i] = rand.Int()  
14    }  
15 }
```

`generateStack()` 函数创建了一个大小为 8192 的 `int` 类型切片。该切片恰好小于等于 64 KB（在 64 位机器环境下，`int` 类型占用 8 字节存储空间），此内存大小计算未包含切片内部字段所占用的内存。  
`generateHeap()` 函数创建了一个大小为 8193 的 `int` 类型切片，该切片恰好大于 64 KB 的内存空间。

```
1 > go build -gcflags="-m" ./cmd/server/main.go
2 # command-line-arguments
3 cmd/server/main.go:6:14: make([]int, 8192) does not escape
4 cmd/server/main.go:13:14: make([]int, 8193) escapes to heap
```

## 动态分配长度不确定时逃逸当长度

`n` 编译时无法确定，Go 编译器无法保证安全地使用栈空间，因此会选择堆分配。

```
1 func generate(n int) {
2     // 创建一个大小不确定的整数切片，大小由参数n决定
3     nums := make([]int, n)
4     for i := 0; i < n; i++ {
5         nums[i] = rand.Int()
6     }
7 }
```

### 分析结果

```
1 cmd/server/main.go:20:14: make([]int, n) escapes to heap
```

`generate(n)` 函数，其创建的切片大小不固定，在调用该函数时需传入相应的大小参数，对象占用内存将在堆上分配。

## 闭包

当闭包引用外部变量时，变量生命周期需延长至闭包执行结束。

```

1 func Increase() func() int {
2     n := 0
3     return func() int {
4         n++
5         return n
6     }
7 }
8
9 func main() {
10     in := Increase()
11     fmt.Println(in())
12 }

```

Increase() 返回值是一个闭包函数，该闭包函数访问了外部变量 n，那变量 n 将会一直存在，直到 in 被销毁。很显然，变量 n 占用的内存不能随着函数 Increase() 的退出而回收，因此将会逃逸到堆上。

```

1 ./escape.go:2:2: moved to heap: n

```

## 实际使用优化

一些响应要求高 流量特别大的项目需要极致性能的项目（比如rta 百万qps）

使用 pprof 进行内存分配分析

<div> <b>pprof</b> VIEW ▾ SAMPLE ▾ REFINE ▾ CONFIG ▾ DOWNLOAD </div>					
Flat	Flat%	Sum%	Cum	Cum%	Name
505.06MB	87.67%	87.67%	576.06MB	100.00%	main.main
71MB	12.33%	100.00%	71MB	12.33%	main.makeSmall
0	0.00%	100.00%	576.06MB	100.00%	runtime.main

main.makeSmall

/Users/yishi/Library/Mobile Documents/com~apple~CloudDocs/工作/2025/浅谈Go逃逸分析及编码建议/main\_gc\_escape.go

Total:	71MB	71MB (flat, cum)	12.33%
22	.	.	// func makeBig() interface{} {
23	.	.	//   obj := BigStruct{}
24	.	.	//   return &obj
25	.	.	// }
26	.	.	// func makeSmall() interface{} {
27	71MB	71MB	//   return &Small{1, 2}
28	.	.	// }
29	.	.	
30	.	.	// var sink []interface{}
31	.	.	
32	.	.	// func main() {

main.main

/Users/yishi/Library/Mobile Documents/com~apple~CloudDocs/工作/2025/浅谈Go逃逸分析及编码建议/main\_gc\_escape.go

Total:	576.06MB	576.06MB (flat, cum)	100%
48	.	.	// go printMemStats("逃逸版本")
49	.	.	
50	.	.	// // 持续产生堆分配
51	.	.	// for {
52	.	.	//     for i := 0; i < 200; i++ {
53	576.06MB	576.06MB	//         sink = append(sink, makeSmall())
54	.	.	//     }
55	.	.	//     time.Sleep(1 * time.Millisecond)
56	.	.	// }
57	.	.	// }
58	.	.	

调整后

pprofVIEWSAMPLEREFINECONFIGDOWNLOAD					
Flat	Flat%	Sum%	Cum	Cum%	Name
67.87MB	100.00%	100.00%	67.87MB	100.00%	main.main
0	0.00%	100.00%	67.87MB	100.00%	runtime.main

## main.main

/Users/yishi/Library/Mobile Documents/com~apple~CloudDocs/工作/2025/浅谈Go逃逸分析及编码建议/main\_gc\_optimized.go

```
Total:      67.87MB    67.87MB (flat, cum)    100%
 49          .          .
 50          .          .
 51          .          .      // func printMemStats(tag string) {
 52          .          .      //     var m runtime.MemStats
 53          .          .      //     for {
 54          .          .      //         time.Sleep(1 * time.Second)
 55          .          .      //         runtime.ReadMemStats(&m)
 56          .          .      //         fmt.Printf("[%s] GC次数=%d, HeapAlloc=%dKB, HeapObjects=%d\n", tag,
 57          .          .      //     }
 58          .          .      // }
```

## runtime.main

/opt/homebrew/Cellar/go@1.23/1.23.10/libexec/src/runtime/proc.go

```
Total:      0    67.87MB (flat, cum)    100%
267          .          .      // A program compiled with -buildmode=c-archive or c-shared
268          .          .      // has a main, but it is not executed.
269          .          .      return
270          .          .      }
271          .          .      fn := main_main // make an indirect call, as the linker doesn't know the a
272          .          .      fn()
273          .          .      if raceenabled {
274          .          .          runExitHooks(0) // run hooks now, since racefini does not return
275          .          .          racefini()
276          .          .      }
277          .          .      }
```

## 对比数据~

GC次数	HeapAlloc	HeapObjects
4	8000KB	169232
5	19241KB	336844
6	29683KB	505444
7	38618KB	671243
7	47123KB	840654
8	47054KB	1003451
8	71360KB	1171456
8	73998KB	1340258
9	72425KB	1501851
9	75086KB	1672061
9	111622KB	1841470
9	114281KB	2011672
10	110379KB	2172653
10	113033KB	2342457
10	115674KB	2511459
10	118258KB	2676861
10	173858KB	2845264
10	176514KB	3015266
10	179186KB	3186274
10	181852KB	3356876
11	174326KB	3517855
11	176990KB	3688259
11	179627KB	3857061
11	182277KB	4026663
11	184930KB	4196465
11	270328KB	4364268
11	272975KB	4533670
11	275572KB	4699872
11	278225KB	4869674
11	280897KB	5040676
11	283547KB	5210278
12	270217KB	5362255
12	272865KB	5531659
12	275499KB	5700261
12	278162KB	5870663
12	280799KB	6039465
12	283437KB	6208267
12	286093KB	6378269
12	288662KB	6542671
12	420669KB	6712074
12	423316KB	6881476
12	425956KB	7050478
12	428569KB	7217680
12	431206KB	7386482

GC次数	HeapAlloc	HeapObjects
0	1343KB	488
0	2760KB	499
1	1826KB	429
1	2995KB	434
2	2798KB	432
2	4640KB	437
3	4315KB	435
3	4316KB	439
3	7205KB	447
3	7205KB	449
4	6676KB	437
4	6677KB	441
4	11205KB	444
4	11205KB	446
4	11205KB	448
5	10365KB	437
5	10365KB	441
5	10365KB	443
5	10365KB	445
5	17453KB	448
5	17453KB	450
5	17453KB	452
5	17453KB	454
5	17453KB	456
6	16124KB	437
6	16125KB	441
6	16125KB	443
6	16125KB	445
6	16125KB	451
6	27213KB	454
6	27213KB	456
6	27213KB	458
6	27213KB	460
6	27213KB	462
6	27213KB	464
6	27213KB	466
7	25129KB	442
7	25130KB	451
7	25130KB	458
7	25131KB	465
7	25131KB	467
7	25131KB	469

场景	建议
高频执行的热点代码路径	必须修复（如循环内的逃逸、核心算法逻辑）
大对象逃逸	建议修复（如缓存大对象到池中，避免频繁堆分配）
低频或非性能敏感代码	无需修复（如初始化函数、一次性操作）
接口或反射导致的逃逸	权衡利弊（若需灵活设计，可接受少量性能损失）

逃逸不是错误，但我们应当意识到它的成本，尤其在高频调用场景下，任何不必要的逃逸都会造成不必要的 GC 压力。

通过逃逸分析与工具的结合使用，我们可以更合理地控制内存分配，从而提升程序的整体性能与稳定性。