

Malloc

- Memory Allocators
 - Introduction
 - C Memory Allocation API
 - Heaps and sbrk
 - Intro to Allocating
 - Placement Strategies
 - Placement Strategy pros and cons
 - Memory Allocator Tutorial
 - Implementing malloc
 - Alignment and rounding up considerations
 - Implementing free
 - Performance
 - Explicit Free Lists Allocators
 - Case study: Buddy Allocator, an example of a segregated list
 - Further Reading
 - Topics
 - Questions/Exercises

Memory memory everywhere but not a allocation to be made - A really fragmented heap

Introduction

Memory allocation is very important! Allocating and de-allocating heap memory is one of the most common operations in any application. The heap at the system level is contiguous series of addresses that the program can expand or contract and use `sbrk` as its own accord ("Overview of Malloc" #ref-mallocinternals). In POSIX, this is called the system break. We use `sbrk` (https://linux.die.net/man/3/sbrk) to move the system break. Most programs don't interact directly with this call, they use a memory allocation system around it to handle chunking up and keeping track of which memory is allocated and which is freed.

We will mainly be looking into simple allocators. Just know that there are other ways of dividing up memory like with `mmap` (https://linux.die.net/man/2/mmap) other allocation schemes and methods like `jemalloc` (https://linux.die.net/man/3/jemalloc)

C Memory Allocation API

- `malloc(size_t bytes)` is a C library call and is used to reserve a contiguous block of memory that may be uninitialized (Jones #ref-jones2010wg14 P. 348). Unlike stack memory, the memory remains allocated until `free` (https://linux.die.net/man/3/free) is called with the same pointer. If `malloc` (https://linux.die.net/man/3/malloc) fails to find at least that much free space requested or `NULL`. That means that `malloc` can return `NULL` even if there is some spaces. Robust programs should check the return value. If your code assumes `malloc` (https://linux.die.net/man/3/malloc) succeeds, and it does not, the program will likely crash (segfault) when it tries to write to address 0. Also, `malloc` may not zero out memory because of performance – check your code to make sure that you are not using uninitialized values.
- `realloc(void *space, size_t bytes)` allows you to resize an existing memory allocation that was previously allocated on the heap (via `malloc`, `calloc`, or `realloc`) (Jones #ref-jones2010wg14 P. 349). The most common use of `realloc` is to resize memory used to hold an array of values. There are two gotchas with `realloc`. One, a new pointer may be returned. Two, it can fail. A naive but readable version of `realloc` is suggested below with sample usage.

```

void * realloc(void * ptr, size_t newsize) {
    // Simple implementation always reserves more memory
    // and has no error checking
    void *result = malloc(newsize);
    size_t oldsize = ... //(depends on allocator's internal data structure)
    if (ptr) memcpy(result, ptr, newsize < oldsize ? newsize : oldsize);
    free(ptr);
    return result;
}

int main() {
    // 1
    int *array = malloc(sizeof(int) * 2);
    array[0] = 10; array[1] = 20;
    // Oops need a bigger array - so use realloc..
    array = realloc(array, 3 * sizeof(int));
    array[2] = 30;
}

```

realloc

The above code is not robust. If <http://ericniebler.com/2014/01/21/realloc/> fails, then the array is leaked. Robust code checks for the return value and only reassigns the original pointer if not null.

```

int main() {
    // 1
    int *array = malloc(sizeof(int) * 2);
    array[0] = 10; array[1] = 20;
    void *tmp = realloc(array, 3 * sizeof(int));
    if (tmp == NULL) {
        // Nothing I can do here
    } else if (tmp == array) {
        // realloc returned same space
        array[2] = 30;
    } else {
        // realloc returned different space
        array = tmp;
        array[2] = 30;
    }
}

```

- `calloc(size_t nmemb, size_t size)` initializes memory contents to zero and also takes two arguments: the number of items and the size in bytes of each item. An advanced discussion of these limitations is

<http://locklessinc.com/articles/calloc/> `calloc` `memset`
[\(http://locklessinc.com/articles/calloc/\)](http://locklessinc.com/articles/calloc/) use [\(http://ericniebler.com/2014/01/21/realloc/\)](http://ericniebler.com/2014/01/21/realloc/) `malloc`

[\(http://ericniebler.com/2014/01/21/realloc/\)](http://ericniebler.com/2014/01/21/realloc/) cause certain performance considerations are taken into account. Note `calloc(x,y)` is identical to `calloc(y,x)`, but you should follow the conventions of the manual. A naive implementation of `calloc` is below.

```
void *calloc(size_t n, size_t size) {
    size_t total = n * size; // Does not check for overflow!
    void *result = malloc(total);

    if (!result) return NULL;

    // If we're using new memory pages
    // just allocated from the system by calling sbrk
    // then they will be zero so zero-ing out is unnecessary,
    // We will be non-robust and memset either way.
    return memset(result, 0, total);
}
```

free

- <https://en.cppreference.com/w/cpp/memory/free> takes a pointer to the start of a piece of memory and makes it available for use in the subsequent calls to the other allocation functions. This is important because we don't want every process in our address space to take an enormous amount of memory. Once we are done using memory, we stop using it with free. A simple usage is below.

```
int *ptr = malloc(sizeof(*ptr));
do_something(ptr);
free(ptr);
```

If you use a piece of memory after it is freed - that is undefined behavior.

Heaps and sbrk

The heap is part of the process memory, and it does not have a fixed size. Heap memory allocation is performed by the C library

when you call [malloc](https://en.cppreference.com/w/cpp/memory/malloc) [calloc](https://en.cppreference.com/w/cpp/memory/calloc) [realloc](https://en.cppreference.com/w/cpp/memory/realloc) [free](https://en.cppreference.com/w/cpp/memory/free) [sbrk](https://en.cppreference.com/w/cpp/memory/sbrk) (https://en.cppreference.com/w/cpp/memory/sbrk) the heap as your program demands more heap memory. As the heap and stack need to grow, we put them at opposite ends of the address space. Stacks don't grow like a heap, new parts of the stack are allocated for new threads. So for typical architectures the heap will grow upwards and the stack grows downwards.

Nowadays, Modern operating system memory allocators no longer need [sbrk](https://en.cppreference.com/w/cpp/memory/sbrk) (https://en.cppreference.com/w/cpp/memory/sbrk) instead they can request independent regions of virtual memory and maintain multiple memory regions. For example, gibibyte requests may be placed in a different memory region than small allocation requests. However this detail is an unwanted complexity.

Programs don't need to call [brk](https://en.cppreference.com/w/cpp/memory/brk) [sbrk](https://en.cppreference.com/w/cpp/memory/sbrk) (https://en.cppreference.com/w/cpp/memory/brk) (https://en.cppreference.com/w/cpp/memory/sbrk) to be interesting because it tells you where your heap currently ends. Instead programs use [malloc](https://en.cppreference.com/w/cpp/memory/malloc) [calloc](https://en.cppreference.com/w/cpp/memory/calloc) [realloc](https://en.cppreference.com/w/cpp/memory/realloc) [free](https://en.cppreference.com/w/cpp/memory/free) (https://en.cppreference.com/w/cpp/memory/malloc) (https://en.cppreference.com/w/cpp/memory/calloc) (https://en.cppreference.com/w/cpp/memory/realloc) (https://en.cppreference.com/w/cpp/memory/free) (https://en.cppreference.com/w/cpp/memory/sbrk) internal implementation of these functions may call [sbrk](https://en.cppreference.com/w/cpp/memory/sbrk) (https://en.cppreference.com/w/cpp/memory/sbrk) when additional heap memory is required.

```
void *top_of_heap = sbrk(0);
malloc(16384);
void *top_of_heap2 = sbrk(0);
printf("The top of heap went from %p to %p \n", top_of_heap, top_of_heap2);
// Example output: The top of heap went from 0x4000 to 0xa000
```

An important fact we just glossed over earlier is memory that was newly obtained by the operating system must be zeroed out. If the operating system did not zero out contents of physical RAM, it might be possible for one process to learn about the memory of another process that had previously used the memory. This would be a security leak. Unfortunately this means that

[malloc](https://en.cppreference.com/w/cpp/memory/malloc) (https://en.cppreference.com/w/cpp/memory/malloc) for [free](https://en.cppreference.com/w/cpp/memory/free) (https://en.cppreference.com/w/cpp/memory/free) freed is often zero. This is unfortunate because, many programmers mistaken write C programs that assume malloc'd memory will always be zero.

```
char* ptr = malloc(300);
// contents is probably zero because we get brand new memory
// so beginner programs appear to work!
// strcpy(ptr, "Some data"); // work with the data
free(ptr);
// later
char *ptr2 = malloc(300); // Contents might now contain existing data and is probably not zer
o
```

Intro to Allocating

Let's try to write Malloc. Here is a first attempt at it – the naive version.

```
void* malloc(size_t size)
// Ask the system for more bytes by extending the heap space.
// sbrk Returns -1 on failure
void *p = sbrk(size);
if(p == (void *) -1) return NULL; // No space left
return p;
}
void free() { /* Do nothing */ }
```

Above is the simplest implementation of malloc, there are a few drawbacks though.

- System calls are slow compared to library calls. We should reserve a large amount of memory and only occasionally ask for more from the system.
- No reuse of freed memory. Our program never re-uses heap memory - it just keeps asking for a bigger heap.

If this allocator was used in a typical program, the process would quickly exhaust all available memory. Instead we need an allocator that can efficiently use heap space and only ask for more memory when necessary. There are programs that may use this type of allocator. Consider a video game allocating objects to load the next scene. It is considerably faster to do the above and just throw the entire block of memory away than it is to do the following placement strategies.

Placement Strategies

During program execution, memory is allocated and de-allocated, so there will be gap in the heap memory that can be re-used for future memory requests. The memory allocator needs to keep track of which parts of the heap are currently allocated and which are parts are available. Suppose our current heap size is 64K, though not all of it is in use because some earlier malloc'd memory has already been freed by the program. Let's say that our heap looks like the following table.

16KiB	10KiB	1KiB	1KiB	30KiB	4KiB	2KiB
free	allocated	free	allocated	free	allocated	free

If a new malloc request for 2KiB is executed (malloc(2048)), where should ^{malloc} (http://en.cppreference.com/cpp/malloc) set the memory? It could use the last 2KiB hole, which happens to be the perfect size! Or it could split one of the other two free holes. These choices represent different placement strategies. Whichever hole is chosen, the allocator will need to split the hole into two. The newly allocated space, which will be returned to the program and a smaller hole if there is spare space left over. A perfect-fit strategy finds the smallest hole that is of sufficient size (at least 2KiB):

16KiB	10KiB	1KiB	1KiB	30KiB	4KiB	2KiB
free	allocated	free	allocated	free	allocated	HERE!

A worst-fit strategy finds the largest hole that is of sufficient size so break the 30KiB hole into two:

16KiB	10KiB	1KiB	1KiB	2KiB	28KiB	4KiB	2KiB
free	allocated	free	allocated	HERE !	free	allocated	free

A first-fit strategy finds the first available hole that is of sufficient size so break the 16KiB hole into two:

2KiB	14KiB	10KiB	1KiB	1KiB	30KiB	4KiB	2KiB
HERE !	free	allocated	free	allocated	free	allocated	free

To introduce another concept, external fragmentation is that even though we have enough memory in the heap, it may be divided up in a way that we are not able to give the full amount. In the example below, of the 64KiB of heap memory, 17KiB is allocated, and 47KiB is free. However the largest available block is only 30KiB because our available unallocated heap memory is fragmented into smaller pieces.

16KiB	10KiB	1KiB	1KiB	30KiB	4KiB	2KiB
free	allocated	free	allocated	free	allocated	free

Placement Strategy pros and cons

The challenges of writing a heap allocator are

- Need to minimize fragmentation (i.e. maximize memory utilization)
- Need high performance
- Fiddly implementation – lots of pointer manipulation using linked lists and pointer arithmetic.
- Both fragmentation and performance depend on the application allocation profile, which can be evaluated but not predicted and in practice, under-specific usage conditions, a special-purpose allocator can often out-perform a general purpose implementation.
- The allocator doesn't know the program's memory allocation requests in advance. Even if we did, this is the http://en.wikipedia.org/wiki/Knapsack_problem (hard!)

Different strategies affect the fragmentation of heap memory in non-obvious ways, which only are discovered by mathematical analysis or careful simulations under real-world conditions (for example simulating the memory allocation requests of a database or web server).

First we will have a more mathematical, one-shot approach to each of these algorithms (Garey, Graham, and Ullman #ref-Garey:1972:WAM:800152.804907). The paper describes a scenario where you have a certain number of bins, and a certain number of allocations, and you are trying to fit the allocations in as few bins as possible, hence using as little memory as possible. The paper discusses theoretical implications and puts a nice limit on the ratios in the long run between the ideal memory usage and the actual memory usage. For those who are interested, the paper concludes that actual memory usage over ideal memory usage as the number of bins increases – the bins can have any distribution – is about 1.7 for first fit and lower bounded by 1.7 for best fit. The problem with this analysis is that very few real world applications need this type of one-shot allocation. Video game object allocations will typically designate a different subheap for each level and fill up that subheap if they need a quick memory allocation scheme that they can throw away.

In practice, we'll be using the result from a more rigorous survey conducted in 2005 (Wilson et al. #ref-10.1007/3-540-60368-9_19). The paper makes sure to note that memory allocation is a moving target. A good allocation scheme to one program may not be a good allocation scheme to another program. Programs don't uniformly follow the distribution of allocations, so very. The survey talks about all the allocation schemes that we have introduced as well as a few extra ones. Here are some summarized takeaways

1. Best fit may have problems when a block is chosen that is almost the right size, and the remaining space is split so small that a program probably won't use it. A way to get around this could be to set a threshold for splitting. This small splitting isn't observed as frequently under regular work load. In addition, the worst case behavior of best fit is bad, but doesn't usually happen p. 43.

2. The paper also talks about an important distinction of first fit. There are multiple notions of first. First could be ordered in terms of time of free, or it could be ordered through the addresses of the start of the block, or it could be ordered by the time of last free – first being least recently used. The survey didn't go too in depth into the performance of each, but did make a note that address ordered and least recently used lists ended up with better performance than the most recently used first.
3. The paper concludes by first saying that under simulated random (assuming uniform at random) work loads, best fit and first fit do as well. Even in practice, both best and address ordered first fit do about as equally as well with a splitting threshold and coalescing. The reasons why aren't entirely known.

Some additional notes we make

1. Best fit may not require a full scan of the heap. When a block of perfect size or of perfect size within a threshold is found, that can be returned, depending on what edge-case policy you have.
2. Worst fit may not need to scan the entire heap as well. Depending on how you set up data structures, your heap could represent the data structure max-heap and each allocation call simply needs to pop the top off, reheapify, and possibly insert a split memory block. Using Fibonacci heaps, that can be extremely inefficient.
3. First fit needs to have some sort of ordering. Most of the time people will default to linked lists which is a fine choice. There isn't too many improvements you can make with a least recently used and most recently used linked list policy, but with address ordered linked lists you can speed up insertion from $O(n)$ to $O(\log(n))$ by using a randomized skip-list in conjunction with your singly linked list. Basically an insert would use the skip list as shortcuts to find the right place to insert the block and a removal would go through the list as normal.
4. There are many placement strategies that we haven't talked about, one is next-fit which is first fit on the next fit block. This adds a sort-of deterministic randomness. You won't be expected to know this algorithm, just know as you are implementing a memory allocator as part of a machine problem, there are more than these.

Memory Allocator Tutorial

A memory allocator needs to keep track of which bytes are currently allocated and which are available for use. This section introduces the implementation and conceptual details of building an allocator, or the actual code that implements `malloc` and `free` (<https://linux.die.net/man/3/free>) (<https://linux.die.net/man/3/malloc>)

Conceptually we are thinking about creating linked lists and lists of blocks! Please enjoy the following ASCII art. `bt` is short for boundary tag.

```
| I am one block | I am another | Tiny block | | | | | | |
|---|---|---|---|---|---|---|---|---|
| meta | space |bt| meta | space|bt| meta | spc|bt|
|-----|-----|-----|
```

We will have implicit pointers in our next block, meaning that we can get from one block to another using addition. This is in contrast to an explicit `metadata *next` field in our meta block.

```
|-----|
p      |
|-----|-----|-----|
| meta | space |bt| meta | space|bt| meta | spc|bt|
|-----|-----|-----|
```

If `p` is char pointer to the start of our block, to get to the next block would be

```
p + sizeof(meta) + p->space + sizeof(boundary_tag)
```

The actual spacing may be different. The metadata can contain different things. A minimal metadata would simply have the size of the block.

Since we write integers and pointers into memory that we already control, we can later consistently hop from one address to the next. This internal information represents some overhead. Meaning even if we had requested 1024 KiB of contiguous memory from the system, we will not be able to provide all of it to the running program.

Our heap memory is a list of blocks where each block is either allocated or unallocated. Thus there is conceptually a list of free blocks, but it is implicit in the form of block size information that we store as part of each block. Let's think of it in terms of a simple implementation.

```
typedef struct {
    size_t block_size;
    char data[0];
} block;
block *p = sbrk(100);
p->size = 100 - sizeof(*p) - sizeof(boundary_tag);
// Other block allocations
```

We could navigate from one block to the next block just by adding the block's size.

```
p + sizeof(meta) + p->block_size + sizeof(boundary_tag)
```

Make sure to get your casting right! Otherwise you will move an extreme amount of bytes over.

The calling program never sees these values. They are internal to the implementation of the memory allocator. As an example, suppose your allocator is asked to reserve 80 bytes (`malloc(80)`) and requires 8 bytes of internal header data. The allocator would need to find an unallocated space of at least 88 bytes. After updating the heap data it would return a pointer to the block. However, the returned pointer does not point to the start of the block because that's where the internal size data is stored! Instead we would return the start of the block + 8 bytes. In the implementation, remember that pointer arithmetic depends on type. For example, `p += 8` adds 8 `sizeof(p)` , not necessarily 8 bytes!

Implementing malloc

The simplest implementation uses first fit. Start at the first block, assuming it exists, and iterate until a block that represents unallocated space of sufficient size is found, or we've checked all the blocks. If no suitable block is found, it's time to call `sbrk()` again to sufficiently extend the size of the heap. For the purposes of this class, we will try to serve every memory request until the operating system tells us we are going to run out of heap space. Other applications may limit themselves to a certain heap size and cause requests to intermittently fail. In addition, a fast implementation might extend it a significant amount so that we will not need to request more heap memory in the near future.

When a free block is found, it may be larger than the space we need. If so, we will create two entries in our implicit list. The first entry is the allocated block, the second entry is the remaining space. There are ways to do this if you want to keep the overhead small. We recommend first for going with readability.

```
typedef struct {
    size_t block_size;
    int is_free;
    char data[0];
} block;
block *p = sbrk(100);
p->size = 100 - sizeof(*p) - sizeof(boundary_tag);
// Other block allocations
```

If you want certain bits to hold different pieces of information, use bit fields!

```
typedef struct {
    unsigned int block_size : 7;
    unsigned int is_free : 1;
} size_free;

typedef struct {
    size_free info;
    char data[0];
} block;
```

The compiler will handle it for you. After setting up your fields then it becomes simply looping through each of the blocks and checking the appropriate fields

Here is a visual representation of what happens. As for some more ascii art, if we assume that we have a block that looks like this, we want to spit if the allocation is let's say 16 bytes

Check if this block is free

```
|-----|
| meta | 40 bytes ----->|bt|
|-----|
```

Here is what the heap will look like after

```
|-----|-----|
| meta | 16 bytes ---->|bt| meta | 12 bytes -->|bt|
|-----|-----|
```

This is before alignment concerns as well.

Alignment and rounding up considerations

Many architectures expect multi-byte primitives to be aligned to some multiple of 2 (4, 16, etc). For example, it's common to require 4-byte types to be aligned to 4-byte boundaries and 8-byte types on 8-byte boundaries. If multi-byte primitives are not stored on a reasonable boundary for example starting at an odd address then the performance can be significantly impacted because it may require two memory read requests instead of one. On some architectures the penalty is even greater - the

http://en.wikipedia.org/wiki/Bus_error#Unaligned_access

program will crash with a [Bus error \(unaligned access\)](http://en.wikipedia.org/wiki/Bus_error#Unaligned_access) in your architecture classes if there was no memory protection.

`malloc`

As https://en.cppreference.com/w/cpp/string/basic_string_view does not know how the user will use the allocated memory, the pointer returned to the program needs to be aligned for the worst case, which is architecture dependent.

`malloc`

From glibc documentation, the glibc https://www.gnu.org/software/glibc/manual/html_node/Memory-Allocation-and-Paging.html uses the following heuristic ("Virtual Memory Allocation and Paging" #ref-vma_paging)

The block that malloc gives you is guaranteed to be aligned so that it can hold any type of data. On GNU systems, the address is always a multiple of eight on most systems, and a multiple of 16 on 64-bit systems." For example, if you need to calculate how many 16 byte units are required, don't forget to round up.

This is what the math would look like in C.

```
int s = (requested_bytes + tag_overhead_bytes + 15) / 16
```


The additional constant ensures incomplete units are rounded up. Note, real code is more likely to symbol sizes e.g.

<http://www.ibm.com/developerworks/library/pa-dalign/>

(<http://www.ibm.com/developerworks/library/pa-dalign/>)

`sizeof(x) - 1`, rather than coding numerical constant 15.

Another added effect is could be internal fragmentation happens when the block you give them is larger than their allocation size. Let's say that we have a free block of size 16B (not including metadata). If they allocate 7 bytes, you may want to round up to 16B and just return the entire block. This gets very sinister when you implementing coalescing and splitting. If you don't implement either, then you may end up returning a block of size 64B for a 7B allocation! There is a *lot* of overhead for that allocation which is what we are trying to avoid.

Implementing free

free

When ~~is called, we need to re-apply the offset~~ (http://www.ibm.com/developerworks/library/pa-dalign/) to get back to the 'real' start of the block – to where we stored the size information. A naive implementation would simply mark the block as unused. If we are storing the block allocation status in a bit field, then we just need to clear the bit:

```
p->info.is_free = 0;
```

However, we have a bit more work to do. If the current block and the next block (if it exists) are both free we need to coalesce these blocks into a single block. Similarly, we also need to check the previous block, too. If that exists and represents an unallocated memory, then we need to coalesce the blocks into a single large block.

To be able to coalesce a free block with a previous free block we will also need to find the previous block, so we store the block's size at the end of the block, too. These are called "boundary tags" (Knuth #ref-knuth1973art). These are Knuth's solution to the coalescing problem both ways. As the blocks are contiguous, the end of one blocks sits right next to the start of the next block. So the current block (apart from the first one) can look a few bytes further back to lookup the size of the previous block. With this information you can now jump backwards!

Take for example a double coalesce. If we wanted to free the middle block we need to turn the surrounding blocks into one big blocks

```
|-----|-----|-----|
| free | space |bt| meta | space|bt| free |spc|bt|
|-----|-----|-----|
```

Becomes a single free block as below

```
|-----|
| free | space ----->|bt|
|-----|
```

Performance

With the above description it's possible to build a memory allocator. It's main advantage is simplicity - at least simple compared to other allocators! Allocating memory is a worst-case linear time operation – search linked lists for a sufficiently large free block. De-allocation is constant time. No more than 3 blocks will need to be coalesced into a single block, and using a most recently used block scheme only one linked list entry.

Using this allocator it is possible to experiment with different placement strategies. For example, you could start searching from where you last free'd a block, or where you last allocated from. If you do store pointers to blocks, you need to be very careful that they always remain valid Particularly when you malloc, free, calloc, realloc, coalesce, split, etc.

Explicit Free Lists Allocators

Better performance can be achieved by implementing an explicit doubly-linked list of free nodes. In that case, we can immediately traverse to the next free block and the previous free block. This can reduce the search time, because the linked list only includes unallocated blocks. A second advantage is that we now have some control over the ordering of the linked list. For example, when a block is free'd, we could choose to insert it into the beginning of the linked list rather than always between its neighbors. We may update our struct to look like this

```
typedef struct {
    size_t info;
    struct block *next;
    char data[0];
} block;
```

Here is what that would look like along with our implicit linked list

Free list pointers above and below

```
>---++                                ++----->
  \/                                  /\
|-----|-----|-----|
| free | space |bt| meta | space|bt| free | spc|bt|
|-----|-----|-----|
  \/                                  /\
++-----++
```

Again apologies for the ascii art, we are really system programmers not computer graphics people.

Where do we store the pointers of our linked list? A simple trick is to realize that the block itself is not being used and store the next and previous pointers as part of the block, though now you have to ensure that the free blocks are always sufficiently large to hold two pointers. We still need to implement Boundary Tags, so we can correctly free blocks and coalesce them with their two neighbors. Consequently, explicit free lists require more code and complexity. With explicit linked lists a fast and simple 'Find-First' algorithm is used to find the first sufficiently large link. However, since the link order can be modified, this corresponds to different placement strategies. For example if the links are maintained from largest to smallest, then this produces a 'Worst-Fit' placement strategy.

Explicit linked list insertion policy

The newly free'd block can be inserted easily into two possible positions: at the beginning or in address order. Inserting at the beginning creates a LIFO (last-in, first-out) policy. The most recently free'd spaces will be reused. Studies suggest fragmentation is worse than using address order (Wilson et al. #ref-10.1007/3-540-60368-9_19).

Inserting in address order ("Address ordered policy") inserts free'd blocks so that the blocks are visited in increasing address order. This policy required more time to free a block because the boundary tags (size data) must be used to find the next and previous unallocated blocks. However, there is less fragmentation.

Case study: Buddy Allocator, an example of a segregated list

A segregated allocator is one that divides the heap into different areas that are handled by different sub-allocators dependent on the size of the allocation request. Sizes are grouped into powers of two and each size is handled by a different sub-allocator and each size maintains its own free list.

A well known allocator of this type is the buddy allocator (Rangan, Raman, and Ramanujam #ref-rangan1999foundations P. 85). We'll discuss the binary buddy allocator which splits allocation into blocks of size 2^n ; $n = 1, 2, 3, \dots$ times some base unit number of bytes, but others also exist like Fibonacci split where the allocation is rounded up to the next Fibonacci number. The basic concept is simple: If there are no free blocks of size 2^n , go to the next level and steal that block and split it into two. If two neighboring blocks of the same size become unallocated, they can be coalesced back together into a single large block of twice the size.

Buddy allocators are fast because the neighboring blocks to coalesce with can be calculated from the free'd block's address, rather than traversing the size tags. Ultimate performance often requires a small amount of assembler code to use a specialized CPU instruction to find the lowest non-zero bit.

The main disadvantage of the Buddy allocator is that they suffer from *internal fragmentation*, because allocations are rounded up to the nearest block size. For example, a 68-byte allocation will require a 128-byte block.

Further Reading

There are many other allocation schemes. One of three allocators used internally by the Linux Kernel. See <http://man7.org/linux/man-pages/man3/malloc.3.html>

(<http://man7.org/linux/man-pages/man3/malloc.3.html>)

- http://en.wikipedia.org/wiki/SLUB_%28software%29
(http://en.wikipedia.org/wiki/SLUB_%28software%29)
- http://en.wikipedia.org/wiki/Buddy_memory_allocation
(http://en.wikipedia.org/wiki/Buddy_memory_allocation)

Topics

- Best Fit
- Worst Fit
- First Fit
- Buddy Allocator
- Internal Fragmentation
- External Fragmentation
- sbrk
- Natural Alignment
- Boundary Tag
- Coalescing
- Splitting
- Slab Allocation/Memory Pool

Questions/Exercises

- What is Internal Fragmentation? When does it become an issue?
- What is External Fragmentation? When does it become an issue?
- What is a Best Fit placement strategy? How is it with External Fragmentation? Time Complexity?
- What is a Worst Fit placement strategy? Is it any better with External Fragmentation? Time Complexity?
- What is the First Fit Placement strategy? It's a little bit better with Fragmentation, right? Expected Time Complexity?
- Let's say that we are using a buddy allocator with a new slab of 64kb. How does it go about allocating 1.5kb?
- When does the 5 line `sbrk` implementation of malloc have a bug? (http://en.cppreference.com/w/cpp/string/basic/basic_string_view)
- What is natural alignment?
- What is Coalescing/Splitting? How do they increase/decrease fragmentation? When can you coalesce or split?
- How do boundary tags work? How can they be used to coalesce or split?

Garey, M. R., R. L. Graham, and J. D. Ullman. 1972. "Worst-Case Analysis of Memory Allocation Algorithms." In **Proceedings of the Fourth Annual Acm Symposium on Theory of Computing**, 143–50. STOC '72. New York, NY, USA: ACM. <https://doi.org/10.1145/800152.804907>

Jones, Larry. 2010. "WG14 N1539 Committee Draft Iso/lec 9899: 201x." International Standards Organization.

Knuth, D.E. 1973. **The Art of Computer Programming: Fundamental Algorithms**. Addison-Wesley Series in Computer Science and Information Processing, v. 1-2. Addison-Wesley.

<https://books.google.com/books?id=dC05RwAACAAJ>
 (<https://books.google.com/books?id=dC05RwAACAAJ>) .

"Overview of Malloc." 2018. *MallocInternals - Glibc Wiki*. Free Software Foundation.

<https://sourceware.org/glibc/wiki/MallocInternals>
 (<https://sourceware.org/glibc/wiki/MallocInternals>)

Rangan, C.P., V. Raman, and R. Ramanujam. 1999. *Foundations of Software Technology and Theoretical Computer Science: 19th Conference, Chennai, India, December 13-15, 1999 Proceedings*. FOUNDATIONS of Computer software Technology and Theoretical Computer Science. Springer.

<https://books.google.com/books?id=0uHME7EfjQEC>
 (<https://books.google.com/books?id=0uHME7EfjQEC>) .

"Virtual Memory Allocation and Paging." 2001. *The GNU C Library - Virtual Memory Allocation and Paging*.

https://ftp.gnu.org/old-gnu/Manuals/glibc-2.2.3/html_chapter/libc_3.html

(https://ftp.gnu.org/old-gnu/Manuals/glibc-2.2.3/html_chapter/libc_3.html) .

Free Software Foundation.

Wilson, Paul R., Mark S. Johnstone, Michael Neely, and David Boles. 1995. "Dynamic Storage Allocation: A Survey and Critical Review." In *Memory Management*, edited by Henry G. Baler, 1–116. Berlin, Heidelberg: Springer Berlin Heidelberg.

