Understanding GC in JSC From Scratch

2022-06-02 ()

Javascript relies on garbage collection (GC) to reclaim memory. In this post, we will dig a little bit into JSC (the Javascript engine of WebKit (https://webkit.org/))'s garbage collection system.

WebKit's blog post on GC (https://webkit.org/blog/7122/introducing-riptide-webkits-retreating-wavefront-concurrent-garbage-collector/) is a great post that explained the novelties of JSC's GC and also positioned it within the context of various GC schemes in academia and industry. However, as someone with little GC background, I found WebKit's blog post too hard to understand, and also too vague to understand the specific design used by JSC. So this blog post attempts to add in some more details, and aims to be understandable even by someone with little prior background on GC.

The garbage collector in JSC is non-compacting (https://en.wikipedia.org/wiki/Tracing_garbage_collection#Moving_vs._non-moving), generational (https://en.wikipedia.org/wiki/Tracing_garbage_collection#Generational_GC_(ephemeral_GC)) and mostly[1]-concurrent (https://en.wikipedia.org/wiki/Tracing_garbage_collection#Stop-the-world_vs._incremental_vs._concurrent). On top of being concurrent, JSC's GC heavily employs lock-free programming for better performance.

As you can imagine, the design used by JSC is quite complex. So instead of diving into the complex invariants and protocols, we will start with the simplest design, and improve it step by step to converge at JSC's design in the end. This way, we not only understand *why* JSC's design works, but also *how* JSC's design is reached.

But first of all, let's get into some background.

Memory Allocation in JSC

Memory allocator and GC are tightly coupled by nature – the allocator allocates memory to be reclaimed by the GC, and the GC frees memory to be reused by the allocator. In this section, we will briefly introduce JSC's memory allocators.

At the core of the memory allocation scheme in JSC is the data structure BlockDirectory (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/BlockDirectory.h.html#JSC::BlockDirectory)[2]. It implements a fixed-sized allocator, that is, an allocator that only allocates memory chunks of some fixed size `S`. The allocator keeps tracks of a list of fixed-sized (in current code, 16KB) memory pages ("blocks") it owns, and a free list. Each block is divided into cells of size `S`, and has a footer at its end[3], which contains various metadata information needed for GC and allocator, e.g., which cells are free. By aggregating and sharing metadata at the footer, it both saves memory and improves performance of related operations: we will go into details later.

When a `BlockDirectory` needs to make an allocation, it tries to allocate from its free list. If the free list is empty, it tries to iterate through the blocks it owns[4], to see if it can find a block containing free cells (which are marked free by GC). If yes, it scans the block footer metadata (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/MarkedBlock.cpp.html#_ZN3JSC11MarkedBlock6Handle5sweepEPNS_8FreeList to find out all the free cells[5] in this block, and put into the free list. Otherwise, it allocates a new block from the OS[6]. Note that this implies a `BlockDirectory`'s free list only contains cells in one block: this is called `m_currentBlock` in the code, and we will revisit this later.

The `BlockDirectory` is used as the building block to build the memory allocators in JSC. JSC employs three kinds of allocators:

1. CompleteSubspace (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/CompleteSubspace.h.html#32): this is a segregated allocator responsible for allocating small objects (max size about 8KB). Specifically, there is a pre-defined list of exponentially-growing size-classes[7], and one `BlockDirectory` is used to handle allocation for each size class. So to allocate an object, you find the smallest size class large enough to hold the object, and allocate from that size class.
2. PreciseAllocation (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/PreciseAllocation.h.html#JSC::PreciseAllocation): this is used to handle large allocations that cannot be handled by `CompleteSubspace` allocator[8]. It simply relies on the standard (malloc-like) memory allocator, though in JSC a custom malloc implementation called `libpas` is used. The downside is that since `PreciseAllocation` is done on a per-object basis, it cannot aggregate and share metadata information of multiple objects together to save memory and improve performance (as `CompleteSubspace`'s block footer did). Therefore, every `PreciseAllocation` comes with a whopping overhead of a 96-byte GC header (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/PreciseAllocation.h.html#JSC::PreciseAllocation) to store the various metadata information needed for GC for this object (though this overhead is justified since each allocation is already at least 8KB).
3. IsoSubspace (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/IsoSubspace.h.html#JSC::IsoSubspace): each `IsoSubspace` is used to allocate objects of a fixed type with a fixed size. So each `IsoSubspace` simply holds a `BlockDirectory` to do allocation (though JSC also has an optimization for small `IsoSubspace` by making them backed by `PreciseAllocation`[9]). This is mainly a security hardening feature that makes use-after-free-based attacks harder[10].

As you can see, `IsoSubspace` is mostly a simplified `CompleteSubspace`, so we will ignore it for the purpose of this post. `CompleteSubspace` is the one that handles the common case: small allocations, and `PreciseAllocation` is mostly the rare slow path for large allocations.

Generational GC Basics

In JSC's generational GC model, the heap consists of a small "new space" (eden), holding the newly allocated objects, and a large "old space" holding the older objects that have survived one GC cycle. Each GC cycle is either an *eden GC* or a *full GC*. New objects are allocated in the eden. When the eden is full, an eden GC is invoked to garbage-collect the unreachable objects in eden. All the surviving objects in eden are then considered to be in the old space[11]. To reclaim objects in the old space, a full GC is needed.

The effectiveness of the above scheme relies on the so-called "generational hypothesis":

1. Most objects collected by the GC are young objects (died when they are still in eden), so eden GC (which only collects the eden) is sufficient to reclaim most of the memory.
2. Pointers from old space to eden is much rarer than pointers from eden to old space or pointers from eden to eden, so an eden GC's runtime is approximately linear to the size of the eden, as it only needs to start from a small subset of the old space. This implies that the cost of GC can be amortized by the cost of allocation.

Inlined vs. Outlined Metadata: Why?

Practically every GC scheme uses some kind of metadata to track which objects are alive. In this section, we will explain how those metadata are stored in JSC, and the motivation behind such design.

In JSC, every object managed by the GC carries the following metadata:

1. Every object managed by GC inherit the JSCell (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/runtime/JSCell.h.html#JSC::JSCell) class, which contains a 1-byte member `cellState`. This `cellState` is a color marker with two colors: white and black[12].
2. Every object also has two out-of-object metadata bits: `isNew` [13] and `isMarked`. For objects allocated by `PreciseAllocation`, the bits reside in the GC header. For objects allocated by `CompleteSubspace`, the bits reside in the block footer.

This may seem odd at first glance since `isNew` and `isMarked` could have been stored in the unused bits of `cellState`. However, this is intentional.

The inlined metadata `cellState` is easy to access for the mutator thread (the thread executing Javascript code), since it is just a field in the object. However, it has bad memory locality for GC and allocators, which need to quickly traverse through all the metadata of all objects in some block owned by `CompleteSubspace` (which is the common case). Outlined metadata have the opposite performance characteristics: they are more expensive to access for the mutator thread, but since they are aggregated into bitvectors and stored in the block footer of each block, GC and allocators can traverse them really fast.

So JSC keeps both inlined and outlined metadata to get the better of both worlds: the mutator thread's fast path will only concern the inlined `cellState`, while the GC and allocator logic can also take advantage of the memory locality of the outlined bits `isNew` and `isMarked`.

Of course, the cost of this is a more complex design... so we have to unfold it bit by bit.

A Really Naive Stop-the-World Generational GC

Let's start with a really naive design just to understand what is needed. We will design a generational, but stop-the-world (i.e. not incremental or concurrent) GC, with no performance optimizations at all. In this design, the mutator side transfers control to the GC subsystem at a "safe point"[14] to start a GC cycle (eden or full). The GC subsystem performs the GC cycle from the beginning to the end (as a result, the application cannot run during this potentially long period, thus "stop-the-world"), and then transfer control back to the mutator side.

For this purpose, let's temporarily forget about `CompleteSubspace`: it is an optimized version of `PrecisionAllocation` for small allocations, and while it is an important optimization, it's easier to understand the GC algorithm without it.

It turns out that in this design, all we need is one `isMarked` bit. The `isMarked` bit will indicate if the object is reachable at the end of the last GC cycle (and consequently, is in the old space, since any object that survived a GC cycle is in old space). All objects are born with `isMarked = false`.

The GC will use a breadth-first search to scan and mark objects. For full GC, we want to reset all `isMarked` bit to `false` at the beginnning, and do a BFS to scan and mark all objects reachable from GC roots. Then all the unmarked objects are known to be dead. For eden GC, we only want to scan the eden space. Fortunately, all objects in the old space are already marked at the end of the previous GC cycle, so they are naturally ignored by the BFS, so we can simply reuse the same BFS algorithm in full GC. In pseudo-code:

Eden GC preparation phase: no work is needed.

Full GC preparation phase[15]:

```
for (JSCell* obj : heap)
  obj->isMarked = false;
```

Eden/Full GC marking phase:

```
while (!queue.empty()) {
  JSCell* obj = queue.pop();
  obj->ForEachChild([&](JSCell* child) {
    if (!child->isMarked) {
      child->isMarked = true;
      queue.push(child);
    }
  });
}
```

Eden/Full GC collection phase:

```
// One can easily imagine optimization to make eden collection
// traverse only the eden space. We ignore it for simplicity.
for (JSCell* obj : heap)
  if (!obj->isMarked)
    free(obj);
```

But where does the scan start, so that we can scan through every reachable object? For full GC, the answer is clear: we just start the scan from all GC roots (https://en.wikipedia.org/wiki/Tracing_garbage_collection#Reachability_of_an_object)[16]. However, for eden GC, in order to reliably scan through all reachable objects, the situation is slightly more complex:

1. Of course, we still need to push the GC roots to the initial queue.
2. If an object in the old space contains a pointer to an object in eden, we need to put the old space object to the initial queue[17].

The invariant for the second case is maintained by the mutator side. Specifically, whenever one writes a pointer slot of some object `A` in the heap to point to another object `B`, one needs to check if `A.isMarked` is `true` and `B.isMarked` is `false`. If so, one needs to put `A` into a "remembered set". Eden GC must treat the objects in the remembered set as if they were GC roots. This is called a `WriteBarrier`. In pseudo-code:

```
// Executed after writing a pointer to 'dst' into a field of 'obj'
if (obj->isMarked && !dst->isMarked)
  addToRememberedSet(obj);
```

Getting Incremental

The stop-the-world GC isn't feasible for production use. A GC cycle (especially a full GC cycle) can take a long time. Since the mutator (application logic) cannot run during the period, the application would appear irresponsive to the user, which is very bad user experience.

A natural way to shorten this irresponsive period is to run GC incrementally: at safe points, the mutator transfers control to the GC. The GC only runs for a short time, doing a portion of the work for the current GC cycle (eden or full), then return control to the mutator. This way, each GC cycle is splitted into many small steps, so the irresponsive periods are less noticeable for the user.

Incremental GC poses a few new challenges to the GC scheme.

The first challenge is the extra interference between GC and mutator: the mutator side, namely the allocator and the `WriteBarrier`, must be prepared to see states arisen from a partially-completed GC cycle. And the GC side must correctly mark all reachable objects despite changes made by the mutator side in between.

Specifically, our full GC must change: imagine that the full GC scanned some object `o` and handed back control to mutator, then the mutator changed a field of `o` to point to some other object `dst`. The object `dst` must not be missed from scanning. Fortunately, in such case `o` will be `isMarked` and `dst` will be `!isMarked` (if `dst` has `isMarked` then it has been scanned, so there's nothing to worry about), so `o` will be put into the remembered set.

Therefore, for full GC to function correctly in the incremental GC scheme, it must consider the remembered set as GC root as well, just like the eden GC.

The other parts of the algorithm as of now can remain unchanged (we leave the proof of correctness as an excerise for the reader). Nevertheless, "what happens if a GC cycle is run partially?" is something that we must keep in mind as we add more optimizations.

The second challenge is that the mutator side can repeatedly put an old space object into the remembered set, and result in redundant work for the GC: for example, the GC popped some object `o` in the remembered set, traversed from it, and handed over control to mutator. The mutator modified `o` again, putting it back to the remembered set. If this happens too often, the incremental GC could do a lot more work than a stop-the-world GC.

The obvious mitigation is to have the GC scan the remembered set last: only when the queue has otherwise been empty do we start popping from the remembered set. However, it turns out that this is not enough. JSC employs a technique called *Space-Time Scheduler* to further mitigate this problem. In short, if it obverves that the mutator was allocating too fast, the mutator would get decreasingly less time quota to run so the GC can catch up (and in the extreme case, the mutator would get zero time quota to run, so it falls back to the stop-the-world approach). The WebKit blog post (https://webkit.org/blog/7122/introducing-riptide-webkits-retreating-wavefront-concurrent-garbage-collector/) has explained it very clearly, so feel free to take a look if you are interested.

Anyway, let's update the pseudo-code for the eden/full GC marking phase:

```
while (!queue.empty() || !rmbSet.empty()) {
  // Both eden GC and full GC needs to consider remembered set
  // Prioritize popping from queue, pop remembered set last
  JSCell* obj = !queue.empty() ? queue.pop() : rmbSet.pop();
  obj->ForEachChild([&](JSCell* child) {
    if (!child->isMarked) {
      child->isMarked = true;
      queue.push(child);
    }
  });
}
```

Incorporate in CompleteSubspace

It's time to get our `CompleteSubspace` allocator back so we don't have to suffer the huge per-object GC header overhead incurred by `PreciseAllocation`.

For `PreciseAllocation`, the actual memory management work is done by `malloc`: when the mutator wants to allocate an object, it just `malloc` it, and when the GC discovers a dead object, it just `free` it.

`CompleteSubspace` introduces another complexity, as it only allocate/deallocate memory from the OS at 16KB-block level, and does memory management itself to divide the blocks into cells that it serves to the application. Therefore, it has to track whether each of its cells is available for allocation. The mutator allocates from the available cells, and the GC marks dead cells as available for allocation again.

The `isMarked` bit is not enough for the `CompleteSubspace` allocator to determine if a cell contains a live object or not: newly allocated objects have `isMarked = false` but are clearly live objects. Therefore, we need another bit.

In fact, there are other good reasons that we need to support checking if a cell contains a live object or not. A canonical example is the conservative stack scanning: JSC cannot precisely understand the layout of the stack, so it needs to treat everything on the stack that could be pointers and pointing to live objects as GC root, and this involves checking if a heap pointer points to a live object or not.

One can easily imagine some kind of `isLive` bit that is `true` for all live objects, which is only flipped to `false` by GC when the object is dead. However, JSC employed a slightly different scheme, which is needed to facilitate optimizations that we will mention later.

As you have seen earlier, the bit used by JSC is called `isNew`.

However, keep in mind: you should **not** think of `isNew` as a bit that tells you **anything** related to its name, or indicates anything by itself. You should think of it as a helper bit, which sole purpose is that, when working togther with `isMarked`, they tell you if a cell contains a live object or not. This thinking mode will be more important in the next section when we introduce logical versioning.

The core invariant around `isNew` and `isMarked` is:

1. At **any** moment, an object is dead iff its `isNew = false` and `isMarked = false`.

If we were a stop-the-world GC, then to maintain this invariant, we only need the following:

1. When an object is born, it has `isNew = true` and `isMarked = false`.
2. At the end of each eden or full GC cycle, we set `isNew` of all objects to `false`.

Then, all newly-allocated objects are live because its `isNew` is `true`. At the end of each GC cycle, an object is live iff its `isMarked` is `true`, so after we set `isNew` to `false` (due to rule 2), the invariant on dead object is maintained, as desired.

However, in an incremental GC, since the state of a partially-run GC cycle can be exposed to mutator, we need to be careful that the invariant holds in this case as well.

Specifically, in full GC, we reset all `isMarked` to `false` at the beginning. Then, during a partially-run GC cycle, the mutator may see a live object with both `isMarked = false` (beacuse it has not been marked by GC yet), and `isNew = false` (because it has survived one prior GC cycle). This violates our invariant.

To fix this, at the beginning of a full GC, we additionally do `isNew |= isMarked` before clearing `isMarked`. Now, during the whole full GC cycle, all live objects must have `isNew = true` [18], so our invariant is maintained. At the end of the cycle, all `isNew` bits are cleared, and as a result, all the unmarked objects become dead, so our invariant is still maintained as desired. So let's update our pseudo-code:

Eden GC preparation phase: no work is needed.

Full GC preparation phase:

```
// Do 'isNew |= isMarked, isMarked = false' for all
// PreciseAllocation and all cells in CompleteSubspace
for (PreciseAllocation* pa : allPreciseAllocations) {
  pa->isNew |= pa->isMarked;
  pa->isMarked = false;
}
for (BlockFooter* block : allCompleteSubspaceBlocks) {
  for (size_t cellId = 0; cellId < block->numCells; cellId++) {
    block->isNew[cellId] |= block->isMarked[cellId];
    block->isMarked[cellId] = false;
  }
}
```

Eden/Full GC collection phase:

```
// Update 'isNew = false' for CompleteSubspace cells
for (BlockFooter* block : allCompleteSubspaceBlocks) {
  for (size_t cellId = 0; cellId < block->numCells; cellId++) {
    block->isNew[cellId] = false;
  }
}
// For PreciseAllocation, in addition to updating 'isNew = false',
// we also need to free the dead objects
for (PreciseAllocation* pa : allPreciseAllocations) {
  pa->isNew = false;
  if (!pa->isMarked)
    free(pa);
}
```

In `CompleteSubspace` allocator, to check if a cell in a block contains a live object (if not, then the cell is available for allocation):

```
bool cellContainsLiveObject(BlockFooter* block, size_t cellId) {
  return block->isMarked[cellId] || block->isNew[cellId];
}
```

Logical Versioning: Do Not Sweep!

We are doing a lot of work at the beginning of a full GC cycle and at the end of any GC cycle, since we have to iterate through all the blocks in `CompleteSubspace` and update their `isMarked` and `isNew` bits. Despite that the bits in one block are clustered into bitvectors thus have good memory locality, this could still be an expensive operation, especially after we have a concurrent GC (as this stage cannot be made concurrent). So we want something better.

The optimization JSC employs is logical versioning. Instead of physically clearing all bits in all blocks for every GC cycle, we only bump a global "logical version", indicating that all the bits are logically cleared (or updated). Only when we actually need to mark a cell in a block during the marking phase do we then physically clear (or update) the bitvectors in this block.

You may ask: why bother with logical versioning, if in the future we still have to update the bitvectors physically anyway? There are two good reasons:

1. If all cells in a block are dead (either died out during this GC cycle[19], or already dead before this GC cycle), then we will never mark anything in the block, so logical versioning enabled us to avoid the work altogether. This also implies that at the end of each GC cycle, it's unnecessary to figure out which blocks become completely empty, as logical versioning makes sure that these empty blocks will not cause overhead to future GC cycles.
2. The marking phase can be done concurrently with multiple threads *and* while the mutator thread is running (our scheme isn't concurrent now, but we will do it soon), while the preparation / collection phase must be performed single-threadedly *and* with the mutator stopped. Therefore, shifting the work to marking phase reduces GC latency in a concurrent setting.

There are two global version number `g_markVersion` and `g_newVersion` [20]. Each block footer also stores its local version number `l_markVersion` and `l_newVersion`.

Let's start with the easier case: the logical versioning for the `isNew` bit.

If you revisit the pseudo-code above, in GC there is only one place where we write `isNew`: at the end of each GC cycle, we set all the `isNew` bits to `false`. Therefore, we simply bump (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/MarkedSpace.cpp.html#446) `g_newVersion` there instead. A local version `l_newVersion` smaller than `g_newVersion` means that all the `isNew` bits in this block have been logically cleared to `false`.

When the `CompleteSubspace` allocator allocates a new object, it needs to start with `isNew = true`. One can clearly do this directly, but JSC did it in a trickier way that involves a block-level bit named `allocated` for slightly better performance. This is not too interesting, so I deferred it to the end of the post, and our scheme described here right now will not employ this optimization (but is otherwise intentionally kept semantically equivalent as JSC):

1. When a `BlockDirectory` starts allocating from a new block, it update the the block's `l_newVersion` to `g_newVersion`, and set `isNew` to `true` for all already-allocated cells (as the block may not be fully empty), and `false` for all available cells.
2. Whenever it allocates a cell, it sets its `isNew` to true.

Why do we want to bother setting `isNew` to `true` for all already-allocated cells in the block? This is to provide a good property. Since we bump `g_newVersion` at the end of every GC cycle, due to the scheme above, for any block with latest `l_newVersion`, a cell is live if and only if its `isNew` bit is set. Now, when checking if a cell is live, if its `l_newVersion` is latest, then we can just return `isNew` without looking at `isMarked`, so our logic is simpler.

The logical versioning for the `isMarked` bit is similar. At the beginning of a full GC cycle, we bump the `g_markVersion` to indicate that all mark bits are logically cleared. Note that the global version is not bumped for eden GC, since eden GC does not clear `isMark` bits.

There is one extra complexity: the above scheme would break down in incremental GC. Specifically, *during* a full GC cycle, we have logically cleared the `isMarked` bit, but we also didn't do anything to the `isNew` bit, so all cells in the old space would appear dead to the allocator. In our old scheme without logical versioning, this case is prevented by doing `isNew |= isMarked` at the start of the full GC, but we cannot do it now with logical versioning.

JSC solves this problem with the following clever trick: *during* a full GC, we should also accept `l_markVersion` that is off-by-one. In that case, we know the `isMarked` bit accurately reflect whether or not a cell is live, since that is the result of the last GC cycle. If you are a bit confused, take a look at footnote[21] for a more elaborated case discussion. It might also help to take a look at the comments in the pseudo-code below:

```
bool cellContainsLiveObject(BlockFooter* block, size_t cellId) {
  if (block->l_newVersion == g_newVersion) {
    // A latest l_newVersion indicates that the cell is live if
    // and only if its 'isNew' bit is set, so we don't need to
    // look at the 'isMarked' bit even if 'isNew' is false
    return block->isNew[cellId];
  }
  // Now we know isNew bit is logically false, so we should
  // look at the isMarked bit to determine if the object is live
  if (isMarkBitLogicallyCleared(block)) {
    // The isMarked bit is logically false
    return false;
  }
  // The isMarked bit is valid and accurately tells us if
  // the object is live or not
  return block->isMarked[cellId];
}

// Return true if the isMarked bitvector is logically cleared
bool isMarkBitLogicallyCleared(BlockFooter* block) {
  if (block->l_markVersion == g_markVersion) {
    // The mark version is up-to-date, so not cleared
    return false;
  }
  if (IsFullGcRunning() && IsGcInMarkingPhase() &&
      block->l_markVersion == g_markVersion - 1) {
    // We are halfway inside a full GC cycle's marking phase,
    // and the mark version is off-by-one, so the isMarked bit
    // should be accepted, and it accurately tells us if the
    // object is live or not
    return false;
  }
  return true;
}
```

Before we mark an object in `CompleteSubspace`, we need to update the `l_markVersion` of the block holding the cell to latest, and materialize the `isMarked` bits of all cells in the block. That is, we need to run the logic at the full GC preparation phase in our old scheme: `isNew |= isMarked`, `isMarked = false` for all cells in the block. This is shown below.

```
// Used by GC marking phase to mark an object in CompleteSubspace
void markObject(BlockFooter* block, size_t cellId) {
  aboutToMark(block);
  block->isMarked[cellId] = true;
}

// Materialize 'isMarked' bits if needed
// To do this, we need to execute the operation at full GC
// prepare phase: isNew |= isMarked, isMarked = false
void aboutToMark(BlockFooter* block) {
  if (block->l_markVersion == g_markVersion) {
    // Our mark version is already up-to-date,
    // which means it has been materialized before
    return;
  }
  // Check if the isMarked bit is logically cleared to false.
  // The function is defined in the previous snippet.
  if (isMarkBitLogicallyCleared(block)) {
    // This means that the isMarked bitvector should
    // be treated as all false. So operation isNew |= isMarked
    // is no-op, so all we need to do is isMarked = false
    for (size_t cellId = 0; cellId < block->numCells; cellId++) {
      block->isMarked[cellId] = false;
    }
  } else {
    // The 'isMarked' bit is not logically cleared. Now let's
    // check if the 'isNew' bit is logically cleared.
    if (block->l_newVersion < g_newVersion) {
      // The isNew bitvector is logically cleared and should be
      // treated as false. So operation isNew |= isMarked becomes
      // isNew = isMarked (note that executing |= is incorrect
      // beacuse isNew could physically contain true!)
      for(size_t cellId = 0; cellId < block->numCells; cellId++) {
        block->isNew[cellId] = block->isMarked[cellId];
        block->isMarked[cellId] = false;
      }
      // We materialized isNew, so update it to latest version
      block->l_newVersion = g_newVersion;
    } else {
      // The l_newVersion is latest, which means that the cell is
      // live if and only if its isNew bit is set.
      // Since isNew already reflects liveness, we do not have to
      // perform the operation isNew |= isMarked (and in fact, it
      // must be a no-op since no dead cell can have isMarked =
      // true). So we only need to do isMarked = false
      for(size_t cellId = 0; cellId < block->numCells; cellId++) {
        block->isMarked[cellId] = false;
      }
    }
  }
  // We finished materializing isMarked, so update the version
  block->l_markVersion = g_markVersion;
}
```

A fun fact: despite that what we conceptually want to do above is `isNew |= isMarked`, the above code never performs a `|=` at all :)

And also, let's update the pseudo-code for relavent GC logic:

Eden GC preparation phase: no work is needed.

Full GC preparation phase:

```
// For PreciseAllocation, we still need to manually do
// 'isNew |= isMarked, isMarked = false' for every allocation
for (PreciseAllocation* pa : allPreciseAllocations) {
  pa->isNew |= pa->isMarked;
  pa->isMarked = false;
}
// For CompleteSubspace, all we need to do is bumping the
// global version for 'isMarked' bit
g_markVersion++;
```

Eden/Full GC collection phase:

```
// For PreciseAllocation, we still need to manually
// update 'isNew = false' for each allocation, and also
// free the object if it is dead
for (PreciseAllocation* pa : allPreciseAllocations) {
  pa->isNew = false;
  if (!pa->isMarked)
    free(pa);
}
// For CompleteSubspace, all we need to do is bumping the
// global version for 'isNew' bit
g_newVersion++;
```

With logical versioning, GC no longer sweeps the `CompleteSubspace` blocks to reclaim dead objects: the reclamation happens lazily, when the allocator starts to allocate from the block. This, however, introduces an unwanted side-effect. Some objects use manual memory management internally: they own additional memory that are not managed by GC, and have C++ destructors to free those memory when the object is dead. This improves performance as it reduces the work of GC. However, now we do not immediately sweep dead objects and run destructor, so the memory that are supposed to be freed by the destructor could be kept around indefinitely longer, if the block is never allocated from. To mitigate this issue, JSC will also periodically sweep the blocks and run the destructors of the dead objects. This is implemented (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/IncrementalSweeper.h.html#JSC::IncrementalSweeper) by `IncrementalSweeper`, but we will not go into details.

To conclude, logical versioning provided two important optimizations to the GC scheme:

1. The so-called "sweep" phase of the GC (to find out and reclaim dead objects) is removed for `CompleteSubspace` objects. The reclamation is done lazily. This is clearly better than sweeping through the block again and again in every GC cycle.
2. The full GC does not need to reset all `isMarked` bit in the preparation phase, but only lazily reset them in the marking phase by `aboutToMark`: this not only reduces work, but also allows the work to be done parallelized and while the mutator is running, after we make our GC scheme concurrent.

Optimizing WriteBarrier: The cellState Bit

As we have explained earlier, whenever the mutator modified a pointer of a marked object `o` to point to an unmarked object, it needs to add `o` to the "remembered set", and this is called the `WriteBarrier`. In this section, we will dig a bit deeper into the `WriteBarrier` and explain the optimizations around it.

The first problem with our current `WriteBarrier` is that the `isMarked` bit resides in the block footer, so retrieving its value requires quite a few computations from the object pointer. Also it doesn't sit in the same CPU cache line as the object, which makes the access even slower. This is undesirable as the cost is paid for every `WriteBarrier`, no matter if we actually added the object to remembered set in the end or not.

The second problem is, our `WriteBarrier` will repeatedly add the same object `o` to the remembered set every time it is run. The obvious solution is to make `rememberedSet` a hash set to de-duplicate the objects it contains, but doing a hash lookup to check if the object already exists is still too expensive.

This is where the last metadata bit that we haven't explained yet: the `cellState` bit comes in, which solves both problems.

Instead of making `rememberedSet` a hash table, we reserve a byte (though we only use 1 bit of it) named `cellState` in every object's object header, to indicate if we might need to put the object into the remembered set in a `WriteBarrier`. Since this bit resides in the object header as an object field (instead of in the block footer), it's trivially accessible to the mutator who has the object pointer.

`cellState` has two possible values: `black` and `white`. The most important two invariants around `cellState` are the following:

1. For any object with `cellState = white`, it is guaranteed that the object does not need to be added to remembered set.
2. Unless *during* a full GC cycle, all `black` (live) objects have `isMarked = true`.

Invariant 1 serves as a fast-path: `WriteBarrier` can return immediately if our object is `white`, and checking it only requires one load instruction (to load `cellState`) and one comparison instruction to validate it is `white`.

However, if the object is `black`, a slow-path is needed to check whether it is actually needed to add the object to remembered set.

Let's look at our new `WriteBarrier`:

```
// Executed after writing a pointer to 'dst' into a field of 'obj'
void WriteBarrier(JSCell* obj) {
  if (obj->cellState == black)
    WriteBarrierSlowPath(obj);
}
```

The first thing to notice is that the `WriteBarrier` is no longer checking if `dst` (the object that the pointer points to) is marked or not. Clearly this does not affect the correctness: we are just making the criteria less restrictive. However, it is unclear to me if we can improve performance while maintaining correctness by making some kind of check on `dst` as well, like the original `WriteBarrier` did.

I wasn't able to get a definite answer on this even from JSC developer. They have two *conjectures* on why they are doing this: first, by not checking `dst`, more objects are put into the remembered set and need to be scanned by GC, so the total amount of work increased. However, the mutator's work probably decreased, as it does less checks and touches less cache lines (by not touching the outlined `isMarked` bit). Of course, the benefit is offsetted by that the mutator is adding more objects into the remembered set, but this isn't too expensive either, as the remembered set is only a segmented vector. GC has to do more work, as it needs to scan and mark more objects. However, after we make our scheme concurrent, the marking phase of GC can be done concurrently as the mutator is running, so the latency is probably[22] hidden. Second, JSC's DFG compiler has optimization pass (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/dfg/DFGStoreBarrierInsertionPhase.cpp.html) that coalesces barriers on the same object together, and the barrier emitted this way naturally cannot check `dst`. Therefore, to make things easier, they simply made all the barriers to not check `dst`. Although these are all conjectures, and it is unclear if adding back the `dst` check can improve performance, this is how JSC works, so let's stick to it.

The interesting part is how the invariants above are maintained by the relavent parties. As always, there are three actors: the mutator (`WriteBarrier`), the allocator, and the GC.

The interaction with the allocator is the simplest. All objects are born `white`. This is correct because newly-born objects are not marked, so have no reason to be remembered.

The interaction with GC is during the GC marking phase:

1. When we mark an object and push it into the queue, we set its `cellState` to `white`.
2. When we pop an object from the queue, before we start to scan its children, we set its `cellState` to `black`.

In pseudo-code, the Eden/Full GC marking phase now looks like the following (Line 5 and Line 9 are the newly-added logic to handle `cellState`, other lines unchanged):

```
1  while (!queue.empty() || !rmbSet.empty()) {
2    // Both eden GC and full GC needs to consider remembered set
3    // Prioritize popping from queue, pop remembered set last
4    JSCell* obj = !queue.empty() ? queue.pop() : rmbSet.pop();
5    obj->cellState = black;       // <---------------- newly added
6    obj->ForEachChild([&](JSCell* child) {
7      if (!child->isMarked) {
8        markObject(child);
9        child->cellState = white; // <---------------- newly added
10       queue.push(child);
11     }
12   });
13 }
```

Let's argue why the invariant is maintained by the above code.

1. For invariant 1, note that in the above code, an object is `white` only if it is inside the queue (as once it's popped out, it becomes `black` again), pending scanning of its children. Therefore, it is guaranteed that the object will still be scanned by the GC later, so we don't need to add the object to remembered set, as desired.
2. For invariant 2, at the end of any GC cycle, any live object is marked, which means it has been scanned, so it is `black`, as desired.

Now let's look at what `WriteBarrierSlowPath` should do. Clearly, it's correct if it simply unconditionally add the object to remembered set, but that also defeats most of the purpose of `cellState` as an optimization mechanism: we want something better.

A top business of `cellState` is to prevent adding an object into the remembered set if it is already there. Therefore, after we put the object into the remembered set, we will set its `cellState` to `white`, like shown below.

```
void WriteBarrierSlowPath(JSCell* obj) {
  obj->cellState = white;
  addToRememberedSet(obj);
}
```

Let's prove why the above code works. Once we added an object to remembered set, we set it to `white`. We don't need to add the same object into the remembered set until it gets popped out from the set by GC. But when GC pops out the object, it would set its `cellState` back to `black`, so we are good.

JSC employed one more optimization. During a full GC, we might see `black` objects that has `isMarked = false` (note that this is the only possible case that the object is unmarked, due to invariant 2). In this case, it's unnecessary to add the object to remembered set, since the object will eventually be scanned in the future (or it becomes dead some time later before it was scanned, in which case we are good as well). Furthermore, we can flip it back to `white`, so we don't have to go into this slow path the next time a `WriteBarrier` on this object runs. To sum up, the optimized version is as below:

```
void WriteBarrierSlowPath(JSCell* obj) {
  if (IsFullGcRunning()) {
    if (!isMarked(obj)) {
      // Do not add the object to remembered set
      // In addition, set cellState to white so this
      // slow path is not triggered on the next run
      obj->cellState = white;
      return;
    }
  } else {
    assert(isMarked(obj));    // due to invariant 2
  }
  obj->cellState = white;
  addToRememberedSet(obj);
}
```

Getting Concurrent and Getting Wild

At this point, we already have a very good incremental and generational garbage collector: the mutator, allocator and GC all have their respective fast-paths for the common cases, and with logical versioning, we avoided redundant work as much as possible. In my humble opinion, this is a good balance point between performance and engineering complexity.

However, obviously, "engineering complexity" is not a word inside JSC's dictionary: after all, they have the most talented engineers, to the point that they even engineered their own purpose-built LLVM from scratch (https://webkit.org/blog/5852/introducing-the-b3-jit-compiler/)!

To squeeze out every bit of performance, JSC proceeded to make the GC scheme concurrent. However, due to the nature of GC, it's often infeasible to use locks to protect against race conditions for performance reasons, so extensive lock-free programming is employed.

But once lock-free programming is involved, one starts to get into all sorts of architecture-dependent memory reordering problems. x86-64 is the more sane architecture: it only requires `StoreLoadFence()`, and it provides somewhat-TSO-like semantics, but JSC also needs ARM64 support for their Apple Sillicon devices. ARM64 has even fewer guarantees: load-load, load-store, store-load, and store-store can all be reordered by the CPU, so any innocent operation could actually need a fence. As if things were not bad enough, for performance reasons, JSC does not want to use too many memory fences on ARM64. So they have the so-called `Dependency` class (https://sillycross.github.io/r/WebKit/Source/WTF/wtf/Atomics.h.html#_ZN3WTF10DependencyC1Ev), which creates an implicit CPU data dependency on ARM64 through some scary assembly hacks, so they can get the desired memory ordering for a specific data-flow without paying the cost of a memory fence. As you can imagine, with all of these complications and optimizations, the code can easily become horrifying.

So due to my limited expertise, it's unsurprising if I missed to explain or mis-explained some important race conditions in the code, especially some ARM64-specific ones: if you spotted any issue in this post, please definitely let me know.

Let's go through the concurrency assumptions first. Javascript is a single-threaded language, so there is always only one mutator thread[23]. Apart from the mutator thread, JSC has a bunch of compilation threads, a GC thread, and a bunch of marking threads. Only the GC marking phase is concurrent: during which the mutator thread, the compiler threads, and a bunch of marking threads are concurrently running (yes, the marking itself is also done in parallel). However, all the other GC phases are run with the mutator thread and compilation threads stopped.

Some Less Interesting Issues

First of all, clearly the `isMarked` and `isNew` bitvector must be made safe for concurrent access, since multiple threads (including marking threads and mutator) may concurrently update it. Using CAS with appropriate retry/bail mechanism is enough for the bitvector itself.

`BlockFooter` is harder, and needs to be protected with a lock: multiple threads could be simutanuously calling `aboutToMark()`, so `aboutToMark()` must be guarded. For the reader side (the `isMarked()` function, which involves first checking if `l_markVersion` is latest, then reading the `isMarked` bitvector), in x86-64 thanks to x86-TSO, one does not need a lock or any memory fence (as long as `aboutToMark` takes care to update `l_markVersion` after the bitvector). In ARM64, since load-load reordering is allowed, a `Dependency` is required.

Making the `cellContainsLiveObject` (or in JSC jargon, `isLive`) check lock-free is harder, since it involves potentially reading both the `isMarked` bit and the `isNew` bit. JSC employs optimistic locking (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/MarkedBlockInlines.h.html#_ZN3JSC11MarkedBlock6Handle6isLiveEjjbPKNS_8H to provide a fast-path. This is not very different from an optimistic locking scheme you can find in a textbook, so I won't dive into the details.

Of course, there are a lot more subtle issues to change. Almost all the pseudo-code above needs to be adapted for concurrency, either by using a lock or CAS, or by using some sort of memory barriers and concurrency protocol to ensure that the code works correctly under concurrency settings. But now let's turn to some more important and tricky issues.

The Race Between WriteBarrier and Marking

One of the most important race is the race between `WriteBarrier` and GC's marking threads. The marking threads and the mutator thread can access the `cellState` of an object concurrently. For performance reasons, a lock is infeasible, so race condition arises.

It's important to note that we call `WriteBarrier` **after** we have written the pointer into the object. This is not only more convenient to use (especially for JIT-generated code), but also allows a few optimizations: for example, in certain cases (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/dfg/DFGStoreBarrierInsertionPhase.cpp.html), multiple writes to the same object may only call `WriteBarrier` once at the end.

With this in mind, let's analyze why our current implementation is buggy. Suppose `o` is an object, and the mutator wants to store a pointer to another object `target` into a field `f` of `o`. The marking logic of GC wants to scan `o` and append its children into the queue. We need to make sure that GC will observe the `o -> target` pointer link.

Let's first look at the correct logic:

Mutator (WriteBarrier)	GC (Marker)
Store(o.f, target)	Store(o.cellState, black)
StoreLoadFence() // WriteBarrier begin	StoreLoadFence()
t1 = Load(o.cellState)	t2 = Load(o.f) // Load a children of o
if (t1 == black): WriteBarrierSlowPath(o)	Do some check to t2 and push it to queue

This is mostly just a copy of the pseudocode in the above sections, except that we have two `StoreLoadFence()`. A `StoreLoadFence()` guarantees th fence may be executed by the CPU out-of-order engine until all `STORE` before the fence have completed. Let's first analyze what could go wrong w fences.

Just to make things perfectly clear, the precondition is `o.cellState = white` (because `o` is in the GC's queue) and `o.f = someOldValue`.

What could go wrong if the mutator `WriteBarrier` doesn't have the fence? Without the fence, the CPU can execute the `LOAD` in line 3 before the `S` the following interleaving:

1. [Mutator Line 3] t1 = Load(o.cellState) // t1 = white
2. [GC Line 1] Store(o.cellState, black)
3. [GC Line 3] t2 = Load(o.f) // t2 = some old value
4. [Mutator Line 1] Store(o.f, target)

Now, the mutator did not add `o` to remembered set (because `t1` is `white`, not `black`), and `t2` in GC is the old value in `o.f` instead of `target`, so `target` into the queue. So the pointer link from `o` to `target` is missed in GC. This can result in `target` being wrongly reclaimed despite it is live.

And what could go wrong if the GC marking logic doesn't have the fence? Similarly, without the fence, the CPU can execute the `LOAD` in line 3 befo Then, in the following interleaving:

1. [GC Line 3] t2 = Load(o.f) // t2 = some old value
2. [Mutator Line 1] Store(o.f, target)
3. [Mutator Line 3] t1 = Load(o.cellState) // t1 = white
4. [GC Line 1] Store(o.cellState, black)

Similar to above, mutator sees `t1 = white` and GC sees `t2 = oldValue`. So `o` is not added to remembered set, and `target` is not pushed into the link is missed.

Finally, let's analyze why the code behaves correctly if both fences are present. Unfortunately there is not a better way than manually enumerating Thanks to the fences, `Mutator Line 1` must execute before `Mutator Line 3`, and `GC Line 1` must execute before `GC Line 3`, but the four lines car reordered arbitrarily. So there are `4! / 2! / 2! = 6` possible interleavings. So let's go!

Interleaving 1:

1. [Mutator Line 1] Store(o.f, target)
2. [Mutator Line 3] t1 = Load(o.cellState) // t1 = white
3. [GC Line 1] Store(o.cellState, black)
4. [GC Line 3] t2 = Load(o.f) // t2 = target

In this interleaving, the mutator did not add `o` to remembered set, but the GC sees `target`, so it's fine.

Interleaving 2:

1. [GC Line 1] Store(o.cellState, black)
2. [GC Line 3] t2 = Load(o.f) // t2 = some old value
3. [Mutator Line 1] Store(o.f, target)
4. [Mutator Line 3] t1 = Load(o.cellState) // t1 = black

In this interleaving, GC saw the old value, but the mutator added `o` to the remembered set, so GC will eventually drain from the remembered set a which time it will see the correct new value `target`, so it's fine.

Interleaving 3:

1. [Mutator Line 1] Store(o.f, target)
2. [GC Line 1] Store(o.cellState, black)
3. [Mutator Line 3] t1 = Load(o.cellState) // t1 = black
4. [GC Line 3] t2 = Load(o.f) // t2 = target

In this interleaving, GC saw the new value `target`, nevertheless, the mutator saw `t1 = black` and added `o` to the remembered set. This is unfortu scan `o` again, but it doesn't affect correctness.

Interleaving 4:

1. [Mutator Line 1] Store(o.f, target)
2. [GC Line 1] Store(o.cellState, black)
3. [GC Line 3] t2 = Load(o.f) // t2 = target
4. [Mutator Line 3] t1 = Load(o.cellState) // t1 = black

Same as Interleaving 3.

Interleaving 5:

1. [GC Line 1] Store(o.cellState, black)
2. [Mutator Line 1] store(o.f, target)
3. [Mutator Line 3] t1 = Load(o.cellState) // t1 = black
4. [GC Line 3] t2 = Load(o.f) // t2 = target

Same as Interleaving 3.

Interleaving 6:

1. [GC Line 1] Store(o.cellState, black)
2. [Mutator Line 1] Store(o.f, target)
3. [GC Line 3] t2 = Load(o.f) // t2 = target
4. [Mutator Line 3] t1 = Load(o.cellState) // t1 = black

Same as Interleaving 3.

This proves that with the two `StoreLoadFence()`, our code is no longer vulnerable to the above race condition.

Another Race Condition Between WriteBarrier and Marking

The above fix alone is not enough: there is another race between `WriteBarrier` and GC marking threads. Recall that in `WriteBarrierSlowPath`, we a object back to `white` if we saw it is not marked (this may happen during a full GC), as illustrated below:

```
... omitted ...
if (!isMarked(obj)) {
  obj->cellState = white;
  return;
}
... omitted ...
```

It turns out that, after setting the object `white`, we need to do a `StoreLoadFence()`, and check again if the object is marked. If it becomes marked, `>cellState` back to `black`.

Without the fix, the code is vulnerable to the following race:

1. [Precondition] `o.cellState = black` and `o.isMarked = false`
2. [WriteBarrier] Check isMarked() // see false
3. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
4. [GC Marking] Popped 'o' from queue, Store(o.cellState, black)
5. [WriteBarrier] Store(o.cellState, white)
6. [Postcondition] `o.cellState = white` and `o.isMarked = true`

The post-condition is bad because `o` will not be added to the remembered set in the future, despite that it needs to be (as the GC has already sca

Let's now prove why the code is correct when the fix is applied. Now the `WriteBarrier` logic looks like this:

1. [WriteBarrier] Store(o.cellState, white)
2. [WriteBarrier] t1 = isMarked()
3. [WriteBarrier] if (t1 == true): Store(o.cellState, black)

Note that we omitted the first "Check isMarked()" line because it must be the first thing executed in the interleaving, as otherwise the `if`-check w

The three lines in `WriteBarrier` cannot be reordered by CPU: Line 1-2 cannot be reordered because of the `StoreLoadFence()`, line 2-3 cannot be r is a store that is only executed if line 2 is true. The two lines in GC cannot be reordered by CPU because line 2 stores to the same field `o.cellStat`

In addition, note that it's fine if at the end of `WriteBarrier`, the object is `black` but GC has only executed to line 1: this is unfortunate, because the this object will add the object to the remembered set despite it's unnecessary. However, it does not affect our correctness. So now, let's enumerat again!

Interleaving 1.

1. [WriteBarrier] Store(o.cellState, white)
2. [WriteBarrier] t1 = isMarked() // t1 = false
3. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // not executed

Object is not marked and white, OK.

Interleaving 2.

1. [WriteBarrier] Store(o.cellState, white)
2. [WriteBarrier] t1 = isMarked() // t1 = false
3. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
4. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // not executed

Object is in queue and white, OK.

Interleaving 3.

1. [WriteBarrier] Store(o.cellState, white)
2. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
3. [WriteBarrier] t1 = isMarked() // t1 = true
4. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // executed

Object is in queue and black, unfortunate but OK.

Interleaving 4.

1. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
2. [WriteBarrier] Store(o.cellState, white)
3. [WriteBarrier] t1 = isMarked() // t1 = true
4. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // executed

Object is in queue and black, unfortunate but OK.

Interleaving 5.

1. [WriteBarrier] Store(o.cellState, white)
2. [WriteBarrier] t1 = isMarked() // t1 = false
3. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
4. [GC Marking] Popped 'o' from queue, Store(o.cellState, black)
5. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // not executed

Object is marked and black, OK.

Interleaving 6.

1. [WriteBarrier] Store(o.cellState, white)
2. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
3. [WriteBarrier] t1 = isMarked() // t1 = true
4. [GC Marking] Popped 'o' from queue, Store(o.cellState, black)
5. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // executed

Object is marked and black, OK.

Interleaving 7.

1. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
2. [WriteBarrier] Store(o.cellState, white)
3. [WriteBarrier] t1 = isMarked() // t1 = true
4. [GC Marking] Popped 'o' from queue, Store(o.cellState, black)
5. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // executed

Object is marked and black, OK.

Interleaving 8.

1. [WriteBarrier] Store(o.cellState, white)
2. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
3. [GC Marking] Popped 'o' from queue, Store(o.cellState, black)
4. [WriteBarrier] t1 = isMarked() // t1 = true
5. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // executed

Object is marked and black, OK.

Interleaving 9.

1. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
2. [WriteBarrier] Store(o.cellState, white)
3. [GC Marking] Popped 'o' from queue, Store(o.cellState, black)
4. [WriteBarrier] t1 = isMarked() // t1 = true
5. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // executed

Object is marked and black, OK.

Interleaving 10.

1. [GC Marking] CAS(o.isMarked, true), Store(o.cellState, white), pushed 'o' into queue
2. [GC Marking] Popped 'o' from queue, Store(o.cellState, black)
3. [WriteBarrier] Store(o.cellState, white)
4. [WriteBarrier] t1 = isMarked() // t1 = true
5. [WriteBarrier] if (t1 == true): Store(o.cellState, black) // executed

Object is marked and black, OK.

So let's update our pseudo-code. However, I would like to note that, in JSC's implementation, they did not use a `StoreLoadFence()` after `obj->cell`
Instead, they made the `obj->cellState = white` a CAS from `black` to `white` (with memory ordering `memory_order_seq_cst`). This is stronger than a
their logic is also correct. Nevertheless, just in case my analysis above missed some other race with other components, our pseudo-code will stick

Mutator `WriteBarrier` pseudo-code:

```
void WriteBarrier(JSCell* obj) {
  StoreLoadFence();            // Note the fence!
  if (obj->cellState == black)
    WriteBarrierSlowPath(obj);
}

void WriteBarrierSlowPath(JSCell* obj) {
  if (IsGcRunning()) {
    if (!isMarked(obj)) {
      if (CompareAndSwap(
         obj->cellState, black /*from*/, white /*to*/) == SUCCESS)
      {
        if (isMarked(obj)) {
          obj->cellState = black;
        }
      }
      return;
    }
  } else {
    assert(isMarked(obj));
  }
  obj->cellState = white;
  // Add 'obj' to remembered set
  rmbSet.push(obj);
}
```

Eden/Full GC Marking phase:

```
while (!queue.empty() || !rmbSet.empty()) {
  JSCell* obj = !queue.empty() ? queue.pop() : rmbSet.pop();
  obj->cellState = black;
  StoreLoadFence();           // Note the fence!
  obj->ForEachChild([&](JSCell* child) {
    if (!child->isMarked) {
      markObject(child);
      child->cellState = white;
      queue.push(child);
    }
  });
}
```

Remove Unnecessary Memory Fence In WriteBarrier

The `WriteBarrier` is now free of hazardous race conditions. However, we are executing a `StoreLoadFence()` for every `WriteBarrier`, which is a ver instruction. Can we optimize it?

The idea is the following: the fence is used to protect against race with GC. Therefore, we definitely need the fence if the GC is concurrently runni is unnecessary if the GC is not running. Therefore, we can check if the GC is running first, and only execute the fence if the GC is indeed running.

JSC is even smarter: instead of having two checks (one that checks if the GC is running and one that checks if the `cellState` is `black`), it combin check for the fast-path where the GC is not running and the object is `white`. The trick is the following:

1. Assume `black = 0` and `white = 1` in the `cellState` enum.
2. Create a global variable called `blackThreshold`. This `blackThreshold` is normally `0`, but at the beginning of a GC cycle, it will be set to `1`, an to `0` at the end of the GC cycle.
3. Now, check if `obj->cellState > blackThreshold`.

Then, if the check succeeded, we know we can immediately return: the only case this check can succeed is when the GC is not running and we ar `blackThreshold = 0` and `cellState = 1` is the only situation to pass the check). This way, the fast path only executes one check. If the check fails, the slow path, which performs the full procedure: check if GC is running, execute a fence if needed, then check if `cellState` is `black` again. In ps

```
void WriteBarrier(JSCell* obj) {
  if (obj->cellState > g_blackThreshold) {
    // Fast-path: the only way to reach here is when
    // the GC is not running and the cellState is white
    return;
  }
  if (!IsGcRunning()) {
    // g_blackThreshold is 0, so our object is
    // actually black, we need to go to WriteBarrierSlowPath
    WriteBarrierSlowPath(obj);
  } else {
    // GC is running so we need to execute the fence
    // and check cellState again
    StoreLoadFence();
    if (obj->cellState == black) {
      WriteBarrierSlowPath(obj);
    }
  }
}
```

Note that there is no race between `WriteBarrier` and GC setting/clearing `IsGcRunning()` flag and changing the `g_blackThreshold` value, because t stopped at a safe point (of course, halfway inside `WriteBarrier` is not a safe point) when the GC starts/finishes.

"Obstruction-Free Double Collect Snapshot"

Concurrent GC also introduced new complexities for the `ForEachChild` function used by GC marking phase to scan all objects referenced by a cer Javascript object has a `Structure` (aka, hidden class) that describes how the content of this object shall be interpreted into object fields. Since th run concurrently with the mutator, and the mutator may change the `Structure` of the object, and may even change the size of the object's butterf that despite the race conditions, it will never crash by dereferencing invalid pointers and never miss to scan a child. Using a lock is clearly infeasib reasons. JSC uses a so-called *obstruction-free double collect snapshot* to solve this problem. Please refer to the Webkit GC blog post (https://webkit.org/blog/7122/introducing-riptide-webkits-retreating-wavefront-concurrent-garbage-collector/) to see how it works.

Some Minor Design Details and Optimizations

You might find this section helpful if you want to actually read and understand the code of JSC, but otherwise feel free to skip it: these details are design, and are not particularly interesting either. I mention them only to bridge the gap between the GC scheme explained in this post and the act JSC.

As explained earlier, each `CompleteSubspace` owns a list of `BlockDirectory` to handle allocations of different sizes; each `BlockDirectory` has an act `m_currentBlock` where it allocates from, and it achieves this by holding a free list of all available cells in the block. But how does it work exactly?

As it turns out, each `BlockDirectory` has a `cursor`, which is reset (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/LocalAllocator.cpp.html#_ZN3JSC14LocalAllocator5resetEv) to point at the begi at the end of an eden or full GC cycle. Until it is reset, it can only move forward. The `BlockDirectory` will move the cursor forward (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/BlockDirectory.cpp.html#_ZN3JSC14BlockDirectory22findBlockForAllocationER until it finds a block containing available cells, and allocate from it. If the cursor reaches the end of the list, it will attempt to steal a 16KB block (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/LocalAllocator.cpp.html#195) from another `BlockDirectory` and allocate from it. will allocate a new 16KB block from OS and allocate from it.

I also mentioned that a `BlockDirectory` uses a free list to allocate from the currently active block `m_currentBlock`. It's important to note that in the implementation of JSC, the cells in `m_currentBlock` does not respect the rule for `isNew` bit. Therefore, to check liveness, one either need to do a s see if the cell is from `m_currentBlock` (for example, see HeapCell::isLive (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/HeapCell.cpp.html#_ZN3JSC8HeapCell6isLiveEv)), or, for the GC[24], stop the m free list (and populate `isNew` in the process), do whatever inspection, then rebuild the free list and resume the mutator. The latter is implemented (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/MarkedBlock.cpp.html#_ZN3JSC11MarkedBlock6Handle14stopAllocatingERKNS (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/MarkedBlock.cpp.html#_ZN3JSC11MarkedBlock6Handle16resumeAllocatingER `stopAllocating()` and `resumeAllocating()`, which are automatically called whenever the world is stopped (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/Heap.cpp.html#_ZN3JSC4Heap16stopThePeripheryENS_11GCConductorE) or re (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/Heap.cpp.html#_ZN3JSC4Heap18resumeThePeripheryEv).

The motivation of allowing `m_currentBlock` to not respect the rule for `isNew` is (a tiny bit of) performance. Instead of manually setting `isNew` to `tr` allocation, a block-level bit `allocated` (aggregated as a bitvector in `BlockDirectory`) is used to indicate if a block is full of live objects. When the f empty (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/MarkedBlock.cpp.html#_ZN3JSC11MarkedBlock6Handle18didConsumeF block is fully allocated), we simply set `allocated` to `true` for this block. When querying cell liveness, we check this bit first

(https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/MarkedBlockInlines.h.html#101) and directly return true if it is set. The `allocate` at the end of each GC cycle (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/BlockDirectory.cpp.html#252), and since the global `isNew` is also bumped, this effectively clears all the `isNew` bits, just as we desired.

JSC's design also support the so-called *constraint solver*, which allows specification of implicit reference edges (i.e., edge not represented as poi This is mainly used to support Javascript interaction with DOM. This part is not covered in this post.

Weak reference has multiple implementations in JSC. The general (but less efficient) implementation is `WeakImpl`, denoting a weak reference edge managing them is `WeakSet`, and you can see it in every block footer (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/MarkedBlock.h.html#JSC::MarkedBlock::Handle::m_weakSet), and in every Prec header (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/PreciseAllocation.h.html#JSC::PreciseAllocation::m_weakSet). However, more efficient specialized implementations to handle the weak map feature in Javascript. The details are not covered in this post.

In JSC, objects may also have destructors. There are three ways the destructors are run. First, when we begin allocating from a block, destructors run. Second, the `IncrementalSweeper` periodically scans the blocks and runs destructors. Finally, when the VM shuts down, the `lastChanceToFinal` called to ensure that all destructors are run at that time. The details of `lastChanceToFinalize()` are not covered in this post.

JSC employs a conservative approach for pointers on the stack and in registers: the GC uses UNIX signals to suspend the mutator thread, so it ca contents and CPU register values to search for data that looks like pointers. However, it's important to note that UNIX signal is **not** used to suspen mutator: the mutator always **actively** suspends itself at a safe point. This is critical, as otherwise it could be suspended at weird places, for examp `HeapCell::isLive` check after it has read `isNew` but before it has read `isMarked`, and then GC did `isNew |= isMarked, isMarked = false`, and boo only reason to suspend the thread is for the GC to get the CPU register values, including the `SP` register value so the GC knows where the stack e if it's possible to do so in a cooperative manner instead of using costly UNIX signals.

Acknowledgements

I thank Saam Barati from JSC team for his enormous help on this blog post. Of course, any mistakes in this post are mine.

Footnotes

1. Brief stop-the-world pause is still required at the start and end of each GC cycle, and may be intentionally performed if the mutator thread (i. Javascript code) is producing garbage too fast for the GC thread to keep up with. ↩

2. The actual allocation logic is implemented in LocalAllocator (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/LocalAllocator.h.html#JSC::LocalAllocator). Despite that in the code `Block` linked list of `LocalAllocator`, (at time of writing, for the codebase version linked in this blog) the linked list always contains exactly one eleme `BlockDirectory` and `LocalAllocator` is one-to-one and can be viewed as an integrated component. This relationship might change in the fut matter for the purpose of this post anyway. ↩

3. Since the footer resides at the end of a 16KB block, and the block is also 16KB aligned, one can do a simple bit math from any object pointer of the block it resides in. ↩

4. Similar to that per-cell information is aggregated and stored in the block footer, per-block information is aggregated as bitvectors and stored fast lookup. Specifically, two bitvectors `empty` and `canAllocateButNotEmpty` track if a block is empty, or partially empty. The code (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/BlockDirectoryBits.h.html#_M/FOR_EACH_BLOCK_DIRECTORY_BIT) is relat because the bitvectors are layouted in a non-standard way (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/BlockDirectoryBits.h.html#JSC::BlockDirectoryBits::Segment) to make resi conceptually it's just one bitvector for each boolean per-block property. ↩

5. While seemingly straightforward, it is not straightforward at all (as you can see in the code). The free cells are marked free by the GC, and du performance optimization the logic becomes very tricky: we will revisit this later. ↩

6. In fact, it also attempts to steal (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/LocalAllocator.cpp.html#195) blocks from o the OS memory allocator may have some special requirements (https://sillycross.github.io/r/WebKit/WTF/Headers/wtf/Gigacage.h.html) requ we ignore those details for simplicity. ↩

7. In the current implementation, the list of sizes (byte) are 16, 32, 48, 64, 80, then `80 * 1.4 ^ n` for `n >= 1` up to about 8KB. Exponential grow the overhead due to internal fragmentation is at most a fraction (in this case, 40%) of the total allocation size. ↩

8. An interesting implementation detail is that `IsoSubspace` and `CompleteSubspace` always return memory aligned to 16 bytes, but `PreciseAlloca` memory address that has reminder 8 module 16. This allows identifying whether an object is allocated by `PreciseAllocation` with a simple bi

9. JSC has another small optimization here. Sometimes a `IsoSubspace` contains so few objects that it's a waste to hold them using a 16KB mem size of `BlockDirectory`). So the first few memory pages of `IsoSubspace` use the so-called "lower-tier", which are smaller memory pages alloc `PreciseAllocation`. In this post, we will ignore this design detail for simplicity. ↩

10. Memory of an `IsoSubspace` is only used by this `IsoSubspace`, never stolen by other allocators. As a result, a memory address in `IsoSubspace` allocate objects of the same type. So for any type `A` allocated by `IsoSubspace`, even if there is a use-after-free bug on type `A`, it is impossib it, allocate type `B` at the same address, and exploit the bug to trick the VM into interpreting an integer field in `B` controlled by attacker as a p

11. In some GC schemes, an eden object is required to survive two (instead of one) eden GC to be considered in old space. The purpose of such sure that any old space object is at least one eden-GC-gap old. In contrast, in JSC's design, an object created immediately before an eden c considered to be in old space immediately, which then can only be reclaimed via a full GC. The performance difference between the two desi conjecture JSC chose its current design because it's easier to make concurrent. ↩

12. There is one additional color `Grey` in the code (https://sillycross.github.io/r/WebKit/Source/JavaScriptCore/heap/CellState.h.html#JSC::CellS turns out that `White` and `Grey` makes no difference (you can verify it by grepping all use of `cellState` and observe that the only comparison checking if it is `Black`). The comments explaining what the colors mean are also a bit outdated. This is likely a historical artifact. In my opinio clean it up and update the comment, as it can easily cause confusion to readers who intend to understand the design. ↩

13. The bit is actually called `isNewlyAllocated` in the code. We shorten it to `isNew` for convenience in this post. ↩

14. *Safe point* is a terminology in GC. At a *safe point*, the heap and stack is in a coherent state understandable by the GC, so the GC can correctl objects are dead or live. ↩

15. For `PreciseAllocation`, all allocated objects are chained into a linked list, so we can traverse all objects (live or dead) easily. This is not effici optimizations for `CompleteSubspace` later. ↩

16. Keep in mind that while this is true for now, as we add more optimizations to the design, this will no longer be true. ↩

17. Note that we push the old space object into the queue, not the eden object, because this pointer could have been overwritten at the start of the eden object potentially collectable. ↩

18. Also note that all objects dead before this GC cycle, i.e. the free cells of a block in `CompleteSubspace`, still have `isNew = false` and `isMarked`

19. Recall that under generational hypothesis, most objects die young. Therefore, that "all objects in an eden block are found dead during eden (completely plausible. ↩

20. In JSC, the version is stored in a `uint32_t` and they have a bunch of logic to handle the case that it overflows `uint32_t`. In my humble opinio overoptimization that results in very hard-to-test edge cases, especially in a concurrent setting. So we will ignore this complexity: one can ea spending 8 more bytes per block footer to have `uint64_t` version number instead. ↩

21. Note that any number of eden GC cycles may have run between the last full GC cycle and the current full GC cycle, but eden GC does not bu for any object born before the last GC cycle (no matter eden or full), the `isMarked` bit honestly reflect if it is live, and we will accept the bit as be off-by-one. For objects born after the last GC cycle, it must have a latest `isNew` version, so we can know it's alive through `isNew`. In both correctly determines if an object is alive, just as desired. ↩

22. And probably not: first, true sharing and false sharing between GC and mutator can cause slowdown. Second, as we have covered before, JS Scheduler to prevent the mutator from allocating too fast while the GC is running. Specifically, the mutator will be intentionally suspended fo duration. So as long as the GC is running, the mutator suffers from an 30%-or-more "performance tax". ↩

23. The real story is a bit more complicated. JSC actually reuse the same VM for different Javascript scripts. However, at any moment, at most o running. So technically, there are multiple mutually-exclusive mutator threads, but this doesn't affect our GC story. ↩

24. The GC needs to inspect a lot of cells, and its logic is already complex enough, so having one less special-case branch is probably beneficial and performance. ↩

‹ NP70PNP + Ubuntu Tweak Notes (../../../05/31/2022-05-31/)

Bizarre Performance Characteristics of Alder Lake CPU (../../11/2022-06-11/) ›

2023 (../../../../archives/2023/) (2)
2022 (../../../../archives/2022/) (9)
2021 (../../../../archives/2021/) (7)

Recents

Debugging a Bit-Flip Error (../../../../2023/06/11/2023-06-11/)

Building a baseline JIT for Lua automatically (../../../../2023/05/12/2023-05-12/)

Building the fastest Lua interpreter.. automatically! (../../../11/22/2022-11-22/)

Pitfalls of using C++ Global Variable Constructor as a Registration Mechanism (../../../10/02/2022-10-02/)

How to check if a real number is an integer in C++? (../../../07/18/2022-07-18/)