# Safety: off How not to shoot yourself in the foot with C++ atomics

#### **Anthony Williams**

Just Software Solutions Ltd http://www.justsoftwaresolutions.co.uk

25th April 2015

# Safety: off How not to shoot yourself in the foot with C++ atomics





# Safety: off How not to shoot yourself in the foot with C++ atomics

- C++ Atomic types and operations
- Worked examples
- Guidelines

# Aside: Profiling



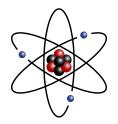
We use atomic operations rather than locks to *improve performance*.

We therefore need to specify the aspect we care about:

- Throughput
- Latency
- Something else

It is vital to profile *before and after* changing to atomic operations

# Atomic types



# Atomic types

- std::atomic<T> provides an atomic type that can store objects of type T.
  - T can be a built in type, or a class type of any size
  - T must be trivially copyable
  - compare\_exchange\_xxx operations require that you can compare T objects with memcmp
  - std::atomic<T> may not be lock free especially for large types
- std::atomic\_flag provides a guaranteed-lock-free flag type.
- The Concurrency TS provides atomic\_shared\_ptr and atomic\_weak\_ptr.

# atomic

**Adjective** 

Meaning

Of or forming a single irreducible unit or component in a larger system.

Origin

Late 15th century: from Old French **atome**, via Latin from Greek **atomos** 'indivisible', based on **a**- 'not' + **temnein** 'to cut'

# Atomic Operations

```
General ops
           load(), store(), exchange(),
           compare_exchange_weak(),
           compare exchange strong()
Arithmetic ops for atomic<Integral> and atomic<T*>
           fetch add().fetch sub()
          ++. --. +=. -=
Bitwise ops for atomic<Integral>
           fetch and(), fetch or(), fetch xor()
          &=, |=, ^=
  Flag ops for atomic_flag
          test and set(), clear()
```

# **Atomic Operations**

```
General ops
                                                                               load(), store(), exchange(),
                                                                               compare_exchange_weak(),
                                                                               compare exchange strong()
Arithmetic apparter at the series of the ser
                                                                               fetch add().fetch sub()
                                                                              ++. --. +=. -=
  Bitwise ops for atomic < Integral >
                                                                               fetch_and(), fetch_or(), fetch_xor()
                                                                               &=, |=, ^=
                  Flag ops for atomic flag
                                                              GUARANTEED LOCK FREE
```

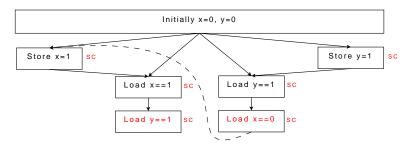
# **Memory Ordering Constraints**

#### 6 values for the ordering on an operation:

- memory\_order\_seq\_cst (the default)
- memory\_order\_acquire
- memory\_order\_release
- memory\_order\_acq\_rel (RMW ops only)
- memory\_order\_relaxed (Experts only)
- memory\_order\_consume (Optimized form of memory\_order\_acquire, for special circumstances, for experts only)

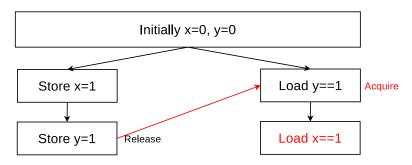
# memory\_order\_seq\_cst ordering

All memory\_order\_seq\_cst operations to all variables form a single total order.



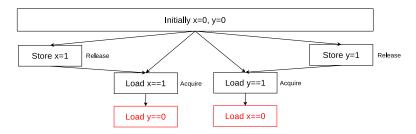
# Release/acquire synchronization

A memory\_order\_release operation synchronizes with a memory\_order\_acquire operation that reads the value written.



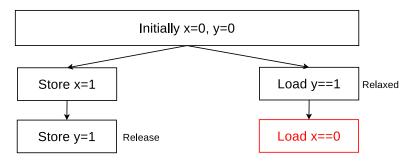
# Release/acquire non-synchronization

Unrelated reads do not synchronize.



# Relaxed atomics: anything can happen

Relaxed atomics can read out of order.





#### C++ has two kinds of fences:

- std::atomic\_thread\_fence⇒ Used for synchronizing between threads
- std::atomic\_signal\_fence
   Used for synchronizing between a thread and a signal handler in that thread

Fences in C++ effectively modify the ordering constraints on neighbouring atomic operations rather than providing any direct ordering constraints themselves.

memory\_order\_acq\_rel fences behave as both
memory\_order\_acquire and memory\_order\_release
fences.

memory\_order\_seq\_cst fences are special: they form part of the total order of memory\_order\_seq\_cst operations, and can therefore enforce orderings beyond the direct pairwise acquire-release orderings. If you're relying on this, you've probably done something wrong.

# Lock-free examples



# Lock-free terminology

#### Obstruction free (Weakest guarantee)

If all other threads are paused then any given thread will complete its operation in a bounded number of steps.

#### Lock free (Most common guarantee)

If multiple threads are operating on a data structure then after a bounded number of steps **one** of them will complete its operation.

#### Wait free (Strongest guarantee)

**Every** thread operating on a data structure will complete its operation in a bounded number of steps, even if other threads are also operating on the data structure.

# Queues



# Why Queues?

- Core facility for communication between threads
- Many types of queue:
  - SPSC / MPSC / MPMC / SPMC
  - bounded / unbounded
  - FIFO / priority / unordered
  - intrusive / non-intrusive
- Good for demonstrating issues

### Lock-based queue

#### Lock-based, unbounded, MPMC, FIFO queue

```
template<typename T>
class queue1{
private:
   std::mutex m;
   std::condition_variable c;
   std::queue<T> q;
};
```

#### Push

```
void push_back(T t) {
    {
       std::lock_guard<std::mutex> guard(m);
       q.push(t);
    }
    c.notify_one();
}
```

# Pop

```
T pop_front() {
   std::unique_lock<std::mutex> guard(m);
   c.wait(guard,[=]{return !q.empty();});
   auto ret=q.front();
   q.pop();
   return ret;
}
```

# Lock-free SPSC FIFO queue

Let's start simple with our lock-free queue:

- One producer thread
- One consumer thread
- Bounded, so no memory allocation
- Assume T has a noexcept copy constructor

# Lock-free SPSC FIFO queue: bounded buffer

```
template<typename T, unsigned buffer_size=42>
class queue2{
  unsigned pop pos{0};
  typedef typename std::aligned_storage<
    sizeof(T),alignof(T)>::type storage_type;
  struct entry{
    std::atomic<bool> initialized{false};
    storage type storage;
  };
  entry buffer[buffer size];
```

# Lock-free SPSC FIFO queue: pushing

```
template<typename T, unsigned buffer_size=42>
class queue2{
  unsigned push pos{0};
public:
  void push_back(T t) {
    unsigned my_pos=push_pos;
    push_pos=(push_pos+1)%buffer_size;
    auto& my_entry=buffer[my_pos];
    while (my_entry.initialized.load()) { }
    new(&my entry.storage) T(t);
    my entry.initialized.store(true);
```

# Aside: avoid busy waits

**Busy waits are to be avoided**: they consume processor power for no purpose.

It is acceptable for a compare\_exchange\_weak loop to have no body: we're hoping to avoid spinning more than a couple of times.

If you need to wait, use a proper wait mechanism such as std::condition\_variable.

# Lock-free SPSC FIFO queue: popping

```
unsigned pop_pos{0};
public:
  T pop front() {
    while(!buffer[pop_pos].initialized.load()){}
    auto ptr=static cast<T*>(
      static cast<void*>(&buffer[pop pos].storage));
    auto ret=*ptr;
    ptr->\sim T();
    buffer[pop_pos].initialized.store(false);
    pop_pos=(pop_pos+1)%buffer size;
    return ret;
```

# Broken Lock-free MPSC FIFO queue

Now let's try and make an MPSC FIFO based on queue2. A naive attempt would be to make push\_pos atomic:

# Broken Lock-free MPSC FIFO queue

Now let's try and make an MPSC FIFO based on queue2. A naive attempt would be to make push\_pos atomic:

This is still broken.

# Broken Lock-free MPSC FIFO queue

- Queue is empty, push\_pos is 0.
- Thread 1 calls push\_back, gets my\_pos is 0, and increments push\_pos to 1.
- Thread 1 checks the cell is empty.
- Thread 1 gets suspended by scheduler
- **Thread 2 calls** push\_back buffer\_size-1 times, so push\_pos loops round to 0.
- Thread 2 calls push\_back again. Thread 2 gets my\_pos of 0, and sets push\_pos to 1.
- Thread 2 checks that the cell is empty.
- Thread 2 populates the cell.
- Thread 1 is woken by the scheduler.
- Thread 1 populates the cell. DATA RACE.

# Not-lock-free MPSC queue

The problem on the previous slide only occurs if the buffer is full. Can we prevent this by checking for a full buffer?

# Not-lock-free MPSC queue

Our queue is now not even obstruction free.

- Queue is empty. push\_pos is 0. pop\_pos is 0.
- 2 Thread 1 calls push\_back and increases size.
- Thread 1 gets my\_pos as 0, increments push\_pos
- Thread 1 is suspended by scheduler.
- Thread 2 pushes buffer\_size-1 entries.
- Thread 2 tries to push another entry, but size==buffer\_size
- Thread 3 calls pop\_front, but pop\_pos is 0 and the entry at 0 hasn't been filled in.
- All threads now stalled waiting for thread 1.

# Fixing queue4

Can we fix this? First we need to identify the problem.

Pushing a value consists of 3 steps:

- Find a free slot in the buffer
- Construct the pushed value in the slot
- Mark the value as available to the consuming thread

# Fixing queue4

Can we fix this? First we need to identify the problem.

Pushing a value consists of 3 steps:

- Find a free slot in the buffer
- Construct the pushed value in the slot
- Mark the value as available to the consuming thread

We need to publish in step 3, rather than step 1.

# Fixing queue4

We need to separate the buffer ordering from the queue ordering, so we need to redo steps 1 and 3.

- Hunt the buffer for a free slot
- Construct the pushed value in the slot
- Link that entry into the queue

Let's use a linked list — that's easy, isn't it? Just push entries on the tail, and pop them off the head.

We still have two locations to update: the next pointer in the previous node, and the tail pointer.

Having the push thread do them in either order can lead to a race.

Answer: update the next pointers from the (one and only) pop thread.

In push\_back we record the previous tail entry:

In pop\_front, if there is no next value for the current entry we can start at the tail and fill them all in:

```
T pop_front() {
  entry* old_head=head;
  while (!old head)
    old head=chase tail();
  head=old head->next;
  auto ptr=static_cast<T*>(
    static cast<void*>(&old head->storage));
  auto ret=*ptr;
  ptr->\sim T();
  recycle node (old head);
  return ret;
```

```
entry* chase tail(){
  entry* next=tail.exchange(nullptr);
  if(!next)
    return nullptr;
  while (next->prev) {
    next->prev->next=next;
    next=next->prev;
  return next;
```

# A lock-free? MPSC FIFO queue

Our queue is now **obstruction free**, but is it **lock-free** or **wait-free**?

- If the queue is full then we have to wait.
  - ⇒ Use a lock-free allocator instead of a fixed buffer.
- If the queue is empty then we have to wait.
- Otherwise, only waiting is in compare-exchange loops
   No upper limit on loops, so cannot be wait-free.
- compare\_exchange\_weak can fail spuriously
   If it does then there is no bound to the number of steps.

Lock-free vs obstruction-free strictly depends on the compare\_exchange\_weak implementation.

# Performance: Cache Ping-Pong



# Performance: Cache Ping-Pong

**Cache Ping-Pong** is where a cacheline is continuously shuttled back and forth between two processors. This occurs when two threads are accessing either:

- the same atomic variable
- different variables on the same cache line

This can have a **big** performance impact, because transferring cache lines is **slow**.

# Cache Ping-Pong in queue5

queue5 can be accessed by many threads in push\_back, and one more thread in pop\_front simultaneously.

```
std::atomic<unsigned> push_hint{0};
entry* head{nullptr};
std::atomic<entry*> tail{nullptr};
entry buffer[buffer_size];
```

head and tail are adjacent, but accessed by different threads  $\Rightarrow$  unnecessary cache ping-pong.

There are many examples in this data structure.

# Cache Ping-Pong avoidance in queue5

The solution to cache ping-pong is to put data on different cache lines by adding padding. **This trades memory space for performance**.

```
std::atomic<unsigned> push_hint{0};
char padding1[padding_size];
entry* head{nullptr};
char padding2[padding_size];
std::atomic<entry*> tail{nullptr};
char padding3[padding_size];
entry buffer[buffer_size];
```

# Cache Ping-Pong avoidance in queue5

Times for 10,000,000 pushes of an integer on each of 3 threads, with another thread popping all 30,000,000 entries.

Run	No padding	With padding	With Lock
1	26.4s	11.4s	27.1s
2	22.4s	9.8s	17.8s
3	22.1s	15.4s	25.4s
4	14.3s	9.0s	24.3s
Mean	21.3s	11.4s	23.7s

# Performance: Memory Ordering Constraints

All the examples so far have used the default ordering constraint: memory\_order\_seq\_cst.

You should use memory\_order\_seq\_cst unless you have a strong reason not to.

# Performance: Memory Ordering Constraints

For x86, only store is affected by the memory order, but for architectures like POWER and ARM with weaker default synchronization, all operations can be affected.

You **must** test on a weakly-ordered system like POWER or ARM if you're using anything other than

memory\_order\_seq\_cst.

# Stacks

### **Stacks**

A stack is a simpler data structure than a queue. It's great for examples, but bad for real use, as all threads are contending to access the top-of-stack.

I'm going to use it to demonstrate a specific problem: the **A-B-A** problem.

# A simple MPSC stack: pushing

```
template<typename T>
class stack1{
  struct node{
    T val;
    node* next;
  };
  std::atomic<node*> head{nullptr};
public:
  void push(T newval){
    auto newnode=new node{newval,head.load()};
    while (!head.compare exchange weak (
            newnode->next, newnode)) {}
```

# A simple MPSC stack: popping

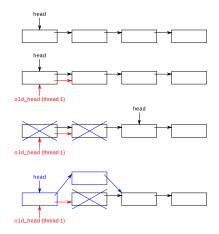
```
I pop() {
  auto old head=head.load();
  for(;;) {
    if (!old head)
      old head=head.load();
    else if (head.compare_exchange_strong (
               old head, old head->next)) {
      auto res=old head->val;
      delete old head;
      return res;
```

# A simple stack: A-B-A issues

Why is this a single-consumer stack? Answer: the A-B-A problem.

# A simple stack: A-B-A issues

- Thread 1 calls pop ()
- Thread 1 reads head into old\_head (A)
- Thread 1 reads old\_head->next
- Thread 1 is suspended
- Thread 2 pops two items, head has new value (B)
- Thread 2 pushes two items
- Second new item is given address of old item, head has original value (A)
- Thread 1 resumes and calls compare\_exchange\_strong, which succeeds because the address is the same
- Stack is now corrupt



### A-B-A

### The setup:

- A value changes from A to B and back to A,
- Other aspects of the data structure have changed, and
- A thread makes a change based on the first time the value was A that is inconsistent with the new state of the data structure.

This most commonly happens where the value is a pointer.

### A-B-A: Solutions

Do not allow a variable to return to its previous value while a thread can do something based on the old value.

Use a change count as part of the variable:

```
struct Value{ T* ptr; unsigned count;};
std::atomic<Value> v;
```

- Ensure that objects are not recycled when still accessible, so A-B-A never happens.
  - ⇒ Reference count the objects, e.g. with std::shared\_ptr and atomic\_shared\_ptr or use hazard pointers, or something similar.

# Reference Counting



# Reference Counting

### How hard can it be?

(Actually, quite hard)

# Basic reference counting

Basic reference counting like std::shared\_ptr is relatively simple.

 $\Rightarrow$  Use an atomic counter, with fetch\_add() and fetch\_sub() to adjust the count.

# Basic reference counting

Basic reference counting like std::shared\_ptr is relatively simple.

 $\Rightarrow$  Use an atomic counter, with fetch\_add() and fetch\_sub() to adjust the count.

This assumes that you have a reference to the count (a std::shared\_ptr instance) which your thread has exclusive access to (for now).

# Basic reference counting

```
struct refcount{
  std::atomic<unsigned> count{1};
  void add ref() {
    count.fetch add(1);
  bool release ref() {
    if (count.fetch_sub(1) == 1) {
      return true;
    return false;
```

# Reference Counting: Memory ordering constraints

We can even put ordering constraints in place:

```
void add_ref() {
  count.fetch_add(1, memory_order_relaxed);
bool release ref() {
  if(count.fetch sub(1, memory order release) == 1) {
    std::atomic thread fence (
      memory order acquire);
    return true;
  return false;
```

# Advanced reference counting

What if we can't guarantee that we have exclusive access to our instance pointer (e.g. a global std::shared\_ptr)?

We need to synchronize somehow and avoid data races and A-B-A issues.

# Advanced reference counting

What if we can't guarantee that we have exclusive access to our instance pointer (e.g. a global std::shared\_ptr)?

We need to synchronize somehow and avoid data races and A-B-A issues.

One way is to count the threads currently accessing each instance.

# Advanced reference counting

Wrap our global reference in a type which holds the pointer to the shared object and the local access count:

```
class atomic_ref_count_ptr{
    struct counted_ptr{
      refcount* p;
      unsigned count;
    };
    std::atomic<counted_ptr> ptr;
public:
    ref_count_ptr load();
    void store(ref_count_ptr newval);
};
```

# Advanced reference counting: load()

- Read the value
- Increase the local count
- Increase the "proper" reference count
- Decrease the local count if we're still pointing to the same object
- If not, release a reference

# Advanced reference counting: load()

```
ref_count_ptr atomic_ref_count_ptr::load() {
  counted ptr local=ptr.load();
  while (local.p && !ptr.compare_exchange_weak (
          local, counted_ptr{local.p, local.count+1})){}
  if(!local.p) return ref_count_ptr(nullptr);
  ref_count_ptr res(local.p);
  counted_ptr newlocal=ptr.load();
  while ((newlocal.p==local.p) &&
        !ptr.compare_exchange_weak(
          newlocal.
          counted_ptr{newlocal.p,newlocal.count-1})){}
  if (newlocal.p!=local.p)
    local.p->release ref();
  return res;
```

# Advanced reference counting: store()

- Add an external reference to the new value
- Store the new value and fetch the old one
- Release the external references on the old one
- If that was the last reference, delete it

# Advanced reference counting: store()

```
void atomic_ref_count_ptr::store(
  ref_count_ptr newval) {
  refcount* newptr=newval.release();
  if(newptr) newptr->add_external();
  counted_ptr old_ptr=ptr.exchange(
    counted_ptr{newptr,0});
  if(old_ptr.p &&
    old_ptr.p->release_external(old_ptr.count))
    delete old_ptr.p;
}
```

# Advanced reference counting: split counters

To handle the external counts we need a new refcount class:

```
struct refcount{
  struct counters{
    unsigned internal;
    unsigned external;
  };
  std::atomic<counters> counts;
  void add ref();
  bool release ref();
  bool release external (unsigned count);
  void add external();
};
```

## Advanced reference counting: split counters

The use of std::atomic<counters> makes the functions a bit more complex: everything is now a CAS loop.

```
void refcount::add ref() {
  counters old=counts.load();
  while (!counts.compare_exchange_weak (
          old,
          counters{old.internal+1,old.external})){}
bool refcount::release_ref() {
  counters old=counts.load();
  while (!counts.compare_exchange_weak (
          old,
          counters{old.internal-1, old.external})){}
  return (!old.external && (old.internal==1));
```

# Advanced reference counting: external counters

```
void refcount::add external(){
  counters old=counts.load();
  while (!counts.compare_exchange_weak (
          old,
          counters{old.internal,old.external+1})){}
bool refcount::release_external(unsigned count) {
  counters old=counts.load();
  while (!counts.compare_exchange_weak (
          old,
          counters{
            old.internal+count-1,old.external-1))){}
  return (old.external==1) &&
    (old.internal == (1-count));
```

# Advanced reference counting: example

Step	Internal count	External count	Local count
One non-atomic ptr	1	0	n/a
Store in atomic ptr	2	1	0
Destroy original non-atomic ptr	1	1	0
One thread starts load()	1	1	1
Second thread starts load()	1	1	2
Another thread calls store () and changes atomic ptr value	1	1	(store-local == 2)
One load thread makes non-atomic ptr	2	1	n/a
Second load thread makes non-atomic ptr	3	1	n/a
One thread finishes load()	2	1	n/a
Second thread finishes load()	1	1	n/a
A			
First non-atomic ptr destroyed	0	1	n/a
Second non-atomic ptr destroyed	-1	1	n/a
Store thread calls release_external() and destroys object	0	0	n/a
В			
First non-atomic ptr destroyed	0	1	n/a
Store thread calls release_external()	1	0	n/a
Second non-atomic ptr destroyed and destroys object	0	0	n/a



- Don't use atomics unless you have to
- Profile before and after
- Test on a weakly-ordered architecture such as POWER or ARM
- Don't use atomics unless you really have to

#### Think in transactions

Do work off to the side and commit with a single atomic operation.

#### Split big operations

If the operation is too big to do in one step, split it into smaller steps that retain the data structure invariants.

#### Limit use cases

Restrict the permitted concurrency levels where possible to reduce implementation complexity.

#### Watch out for ABA problems

These require the circumstances to align just so, but will destroy your data structure when they happen. They can be easily missed in testing.

#### Avoid cache ping pong

Add padding between variables that are accessed from different threads. Try and avoid too many threads accessing the same variable.

#### Stick to memory\_order\_seq\_cst

Unless you **really** know what you're doing, and **really** need the performance gain, stick to the default memory\_order\_seq\_cst. Anything else can be a nightmare to prove correct.

#### Package things up

Wrap atomic operations with types that only expose the desired functionality, to clarify the user code and hide the complexity.

#### Aim for lock-free

Aim for your code to be at least *obstruction-free*, and preferably *lock-free*. Leave *wait-free* for those rare circumstances where you need it.

#### Avoid busy waits

If you're actually waiting (as opposed to spinning on a compare\_exchange\_weak operation), use a proper wait mechanism.

# Questions?

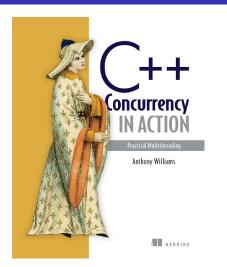
#### Just::Thread



just::thread provides a complete implementation of the C++14 thread library for MSVC and g++ on Windows, and g++ for Linux and MacOSX.

Just::Thread Pro gives you actors, concurrent hash maps, concurrent queues and synchronized values.

## My Book



C++ Concurrency in Action: Practical Multithreading

http://stdthread.com/book

#### Picture credits

The images listed below are from the specified source, with the specified license. All other images are copyright Just Software Solutions Ltd, licensed under Creative Commons Attribution ShareAlike 4 https://creativecommons.org/licenses/by-sa/4.0/.

- Safety Switch: https://www.flickr.com/photos/jamescridland/6163838972/ by James Cridland, Creative Commons Attribution
- Stop watch: https://www.flickr.com/photos/o5com/5488964999/ by o5com, Creative Commons Attribution-NoDerivs
- Lithium Atom: https://commons.wikimedia.org/wiki/File:Stylised\_Lithium\_Atom.svg by Indolences and Rainer Klute, Creative Commons Attribution-ShareAlike
- Fence: http://www.public-domain-image.com/full-image/
  nature-landscapes-public-domain-images-pictures/sunshine-public-domain-images-pictures/
  sunlight-over-picket-fence.jpg-royalty-free-stock-photograph.html by Leon Brooks, Public Domain
- Lock image: http://pixabay.com/en/lock-closed-shut-keyhole-306311/ by Nemo, Public Domain
   "Unauthorised" overlay: http://pixabay.com/en/unauthorised-denied-ban-prohibition-156169/ by OpenClips, Public Domain
- Queue for Apple Store: http://www.geograph.org.uk/photo/3143246 by Robin Stott, Creative Commons
- Ping Pong Set: https://commons.wikimedia.org/wiki/File:Ping-Pong\_2.jpg by Daniel Schwen, Creative Commons Attribution-ShareAlike
- Stack of presents: http://christmasstockimages.com/free/ideas\_concepts/slides/christmas\_gift\_stack.htm by christmasstockimages.com, Creative Commons Attribution
- Abacus: https://commons.wikimedia.org/wiki/File:Abacus\_6.jpg by Loadmaster (David R. Tribble), Creative Commons Attribution ShareAlike
- Success sign: https://commons.wikimedia.org/wiki/File:Success\_sign.jpg by Keith Ramsey (RambergMedialmages), Creative Commons Attribution ShareAlike