How to Write a Custom Allocator

Bob Steagall CppCon 2017



Bob Steagall CppCon 2017

On Writing Allocators

We've all heard heroic tales of other people doing this.

- John McFarlane, CppCon 2017

Overview

- What is an allocator and why write your own?
- Some important background
- Modern allocator facilities
- A container's perspective
- A few guidelines

What is an Allocator?

The basic purpose of an allocator is to provide a source of memory for a given type, and a place to return that memory to once it is no longer needed.

- Bjarne Stroustrup, TCPL, 4th Edition

A service that grants exclusive use of a region of memory to a client.

- Alisdair Meredith, CppCon 2017

What is a Standard Library Allocator?

- A kind of object used by the C++ standard library to manage memory
 - Implementation detail of standard containers, as a template parameter

```
template <class T, class Allocator=allocator<T>> class vector;
```

- Invented by Alexander Stepanov as part of the original STL
 - Attempt to solve near/far/huge pointer types on PCs at the time
 - · Effort to the library more flexible and independent of the underlying memory model
- Containers have special requirements
 - Need an interface that is more "granular" than new and delete
 - Needed finer control

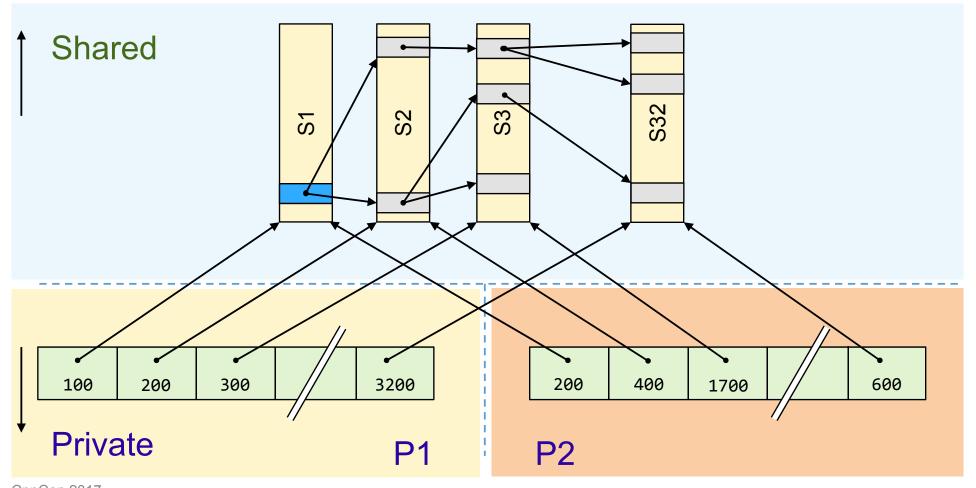
Allocator Mission

- Perform allocation / construction / destruction / deallocation services
 - Separate allocation from construction
 - Separate destruction from deallocation
- Encapsulate information about allocation strategy
- Encapsulate information about addressing model
- Hide memory management and addressing model details from containers
- Support reuse of allocation strategies across container types

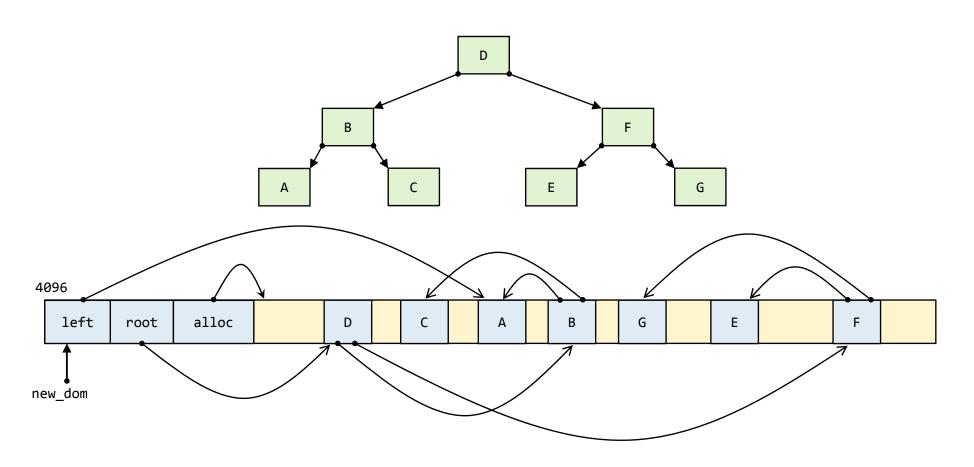
Why Write Your Own Allocator?

- Higher performance
 - Stack-based allocation
 - Per-container private allocation
 - Per-thread allocation
 - Pooled / slab allocation
 - Arena allocation
- Debug / instrumented / test
- Relocatable data
 - Shared memory
 - Self-contained heaps

Motivating Example – Shared Memory Data Structures

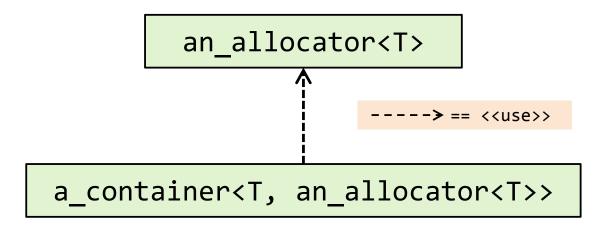


Motivating Example – Self-Contained DOM



A Brief Allocator History

C++03 Allocators



C++03 Allocators

 Containers obtained their allocation services directly from their allocator template argument:

```
T* p = my_alloc.allocate(1u);
```

Containers were free to assume that:

C++03 Default Allocator

```
template<class T> struct allocator
    typedef size t
                       size type;
    typedef ptfdiff t
                       difference type:
    typedef T*
                       pointer:
                       const pointer;
    typedef T const*
    typedef T&
                       reference;
    typedef T const&
                      const reference;
    typedef T
                       value type;
    template<class U> struct rebind { typedef allocator<U> other; };
    pointer allocate(size type n, allocator<void>::const pointer hint = 0);
    void
            deallocate(pointer p, size type n);
    void
            construct(pointer p, T const& val);
    void
            destroy(pointer p);
};
template<class T, class U> bool operator ==(allocator<T> const&, allocator<U> const&);
template<class T, class U> bool operator !=(allocator<T> const&, allocator<U> const&);
```

C++03 Default Allocator – Allocation/Deallocation

```
template<class T> inline T*
allocator<T>::allocate(size_t n)
    return static_cast<T*>(::operator new(n * sizeof(T)));
template<class T> inline void
allocator<T>::deallocate(T* p, size_t)
    ::operator delete(p);
```

C++03 Default Allocator – Construction/Destruction

```
template<class T> inline void
allocator<T>::construct(pointer p, T const& val)
    ::new((void*)p) T(val);
template<class T> inline void
allocator<T>::destroy(pointer p)
    return ((T*)p)->~T();
```

C++03 Default Allocator – Comparison

```
template<class T, class U> inline bool
operator ==(allocator<T> const&, allocator<U> const&)
    return true;
template<class T, class U> inline bool
operator !=(allocator<T> const&, allocator<U> const&)
    return false;
```

C++03 Default Allocator

```
template<class T> struct allocator
    typedef size t
                       size type;
    typedef ptfdiff t
                       difference type:
    typedef T*
                       pointer:
                       const pointer;
    typedef T const*
    typedef T&
                       reference;
    typedef T const&
                       const reference;
    typedef T
                       value type;
    template<class U> struct rebind { typedef allocator<U> other; };
    pointer allocate(size type n, allocator<void>::const pointer hint = 0);
    void
            deallocate(pointer p, size type n);
    void
            construct(pointer p, T const& val);
    void
            destroy(pointer p);
};
template<class T, class U> bool operator ==(allocator<T> const&, allocator<U> const&);
template<class T, class U> bool operator !=(allocator<T> const&, allocator<U> const&);
```

C++03 Default Allocator – Rebinding

```
template<class T, class Alloc = allocator<T>>
class list
    typedef typename Alloc::template rebind<list node<T>>::other node allocator;
    typedef typename node allocator::pointer node pointer;
};
template<class T, class Alloc> typename list<T,Alloc>::node pointer
list<T,Alloc>::alloc node(T const& t)
   node allocator na(this->m_alloc);
   na.construct(np, t);
   return np;
```

C++03 Allocators – Implications

- Implementations assumed that pointers were always T*
 - No support for synthetic pointers / alternate addressing models
- Implementations assumed that instances are always equal and interchangeable
 - Standard containers did not support stateful allocators
- Many interesting problems could not be solved using the standard containers
 - For example: shared memory data structures, self-contained heaps
- Scoped allocation was tricky
 - Consider the case of map<string, vector<list<string>>>

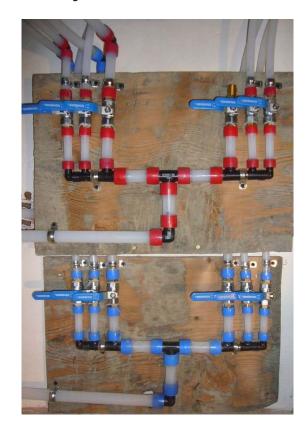
Modern C++ Allocator Facilities

Modern Allocator Plumbing

At first, it may seem like this...

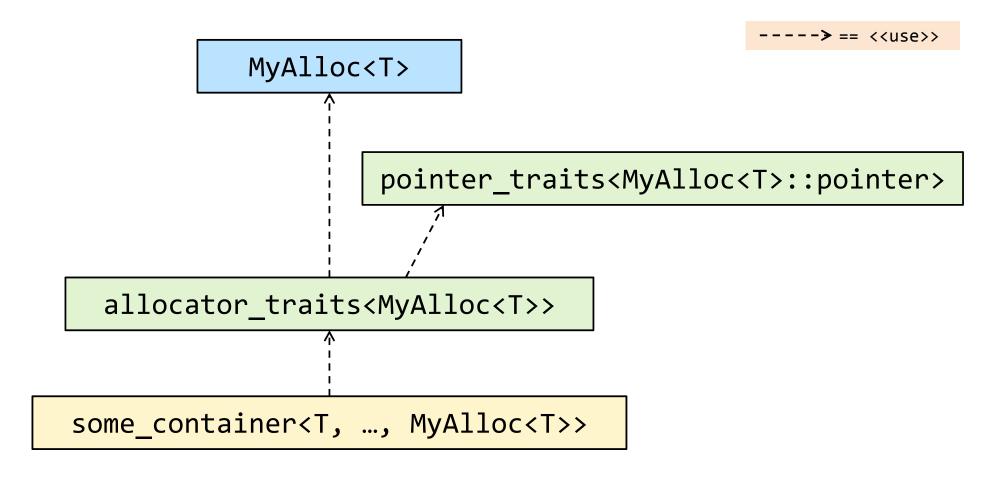


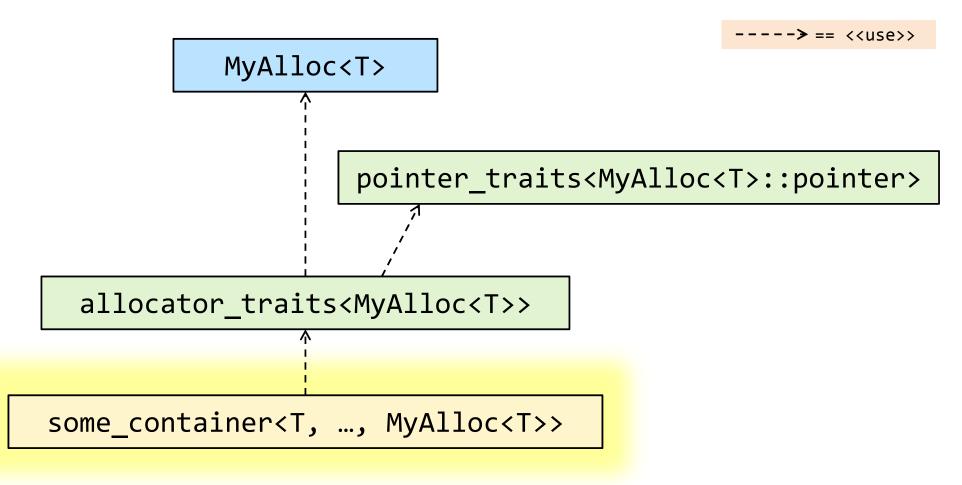
• But really, it's closer to this...

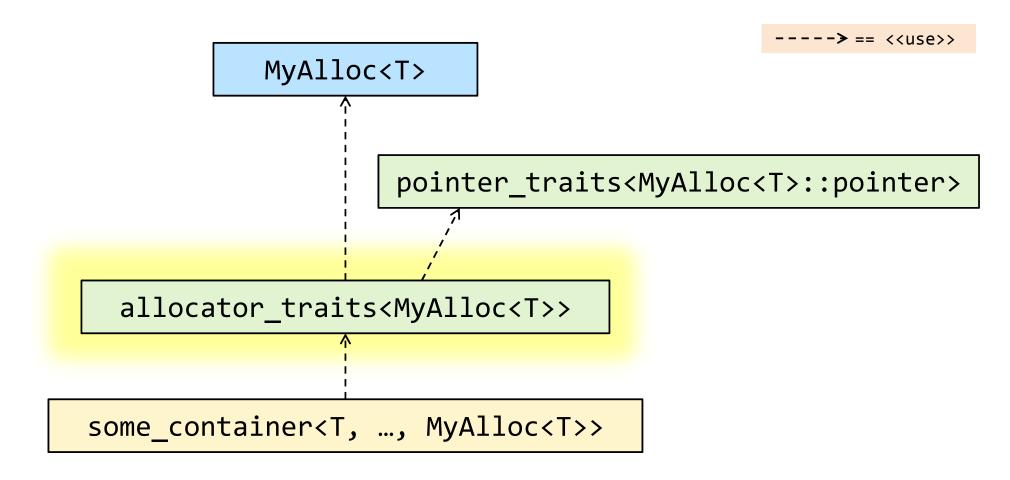


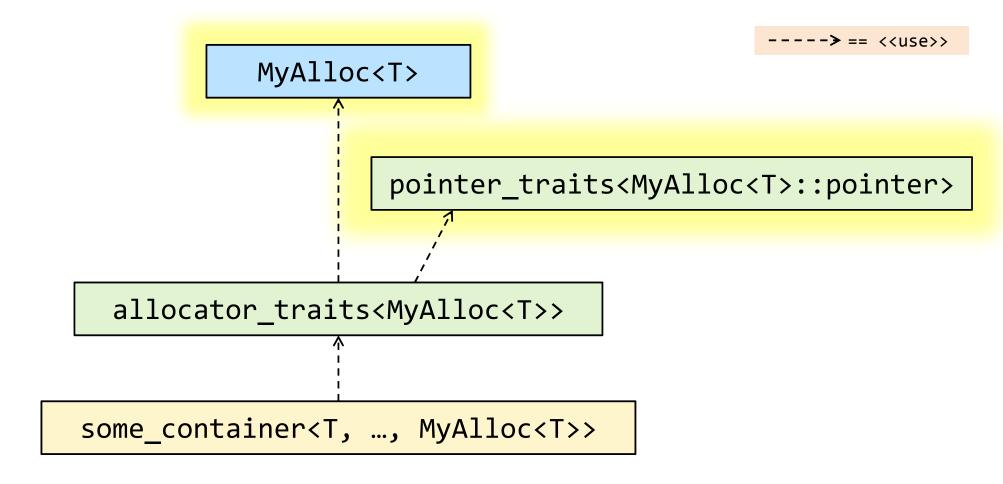
New Requirements to Improve Allocator Support in Modern C++

- nullablepointer.requirements
 - Pointer-like type that supports null values
- allocator.requirements
 - Defines what allocator is and its relationship to allocator traits
- pointer.traits
 - Describes a uniform interface to pointer-like types used by allocator_traits
- allocator traits
 - Describes uniform interface to allocator types used by containers
- allocator.adaptor
 - Describes an adaptor template that supports deep propagation of allocators.
- container.requirements.general
 - Defines <u>allocator-aware container</u> requirements









Allocator Awareness

- What is it?
 - Pedantically fulfilling the requirements in Table 86 (see N4687)
- Allocator-awareness means doing something that makes sense when a container uses a non-default allocator
- In modern C++, allocators
 - Can employ synthetic pointers (aka pointer-like types) for addressing operations
 - Can be stateful
 - Can compare non-equal and be non-interchangeable
 - Can have non-trivial copy/move/swap semantics (propagate)

A Quick Word From Our Sponsors About Allocator Propagation

- Lateral propagation refers to what happens to a container's allocator during
 - Copy/move construction
 - Copy/move assignment
 - Swap
- <u>Deep</u> propagation refers to "nesting" the allocator of the outermost container in a container hierarchy
 - Consider once more map<string, vector<list<string>>>
 - Modern C++ has scoped_allocator_adaptor to help with this situation

Today's focus will be on lateral propagation

Implications of Allocator Awareness for Library Implementors

- Containers must now perform all allocator-related operations using allocator_traits
- Containers must support the case of non-equal allocators
- Containers must support lateral propagation
- Containers now include allocator-extended constructors to support deep propagation
- Containers no longer use reference or const_reference nested type aliases from allocators
- allocator_traits::construct() uses perfect forwarding
- allocator_traits::construct()/destroy() now use T* instead of the pointer type alias

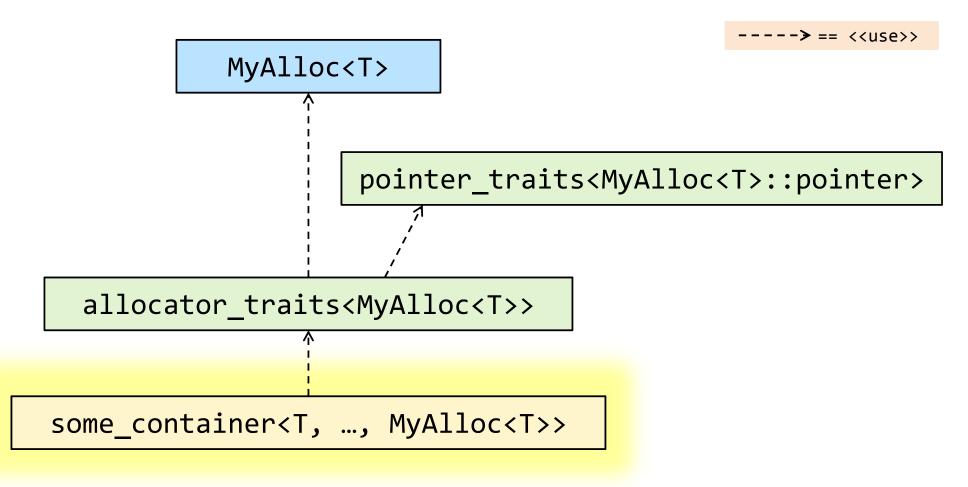
Implications of Allocator Awareness for Library Users

- Stateful allocators can be used with the standard containers
- It is now (relatively) safe and straightforward to use scoped allocators for nested container hierarchies
- Some mutating container operations may lose noexcept property (depending on implementation)

Implications of Allocator Awareness for Allocator Implementors

- Some of the previous verbiage and burden is removed
 - No need to define reference or const_reference typedefs
 - No need to implement construct() or destroy() member functions
 - Only need to specify properties and services that are not the defaults provided by allocator_traits
- However, allocator implementors must now consider
 - How to specify copy/move/swap operations for stateful allocators
 - How to represent pointers for non-traditional addressing

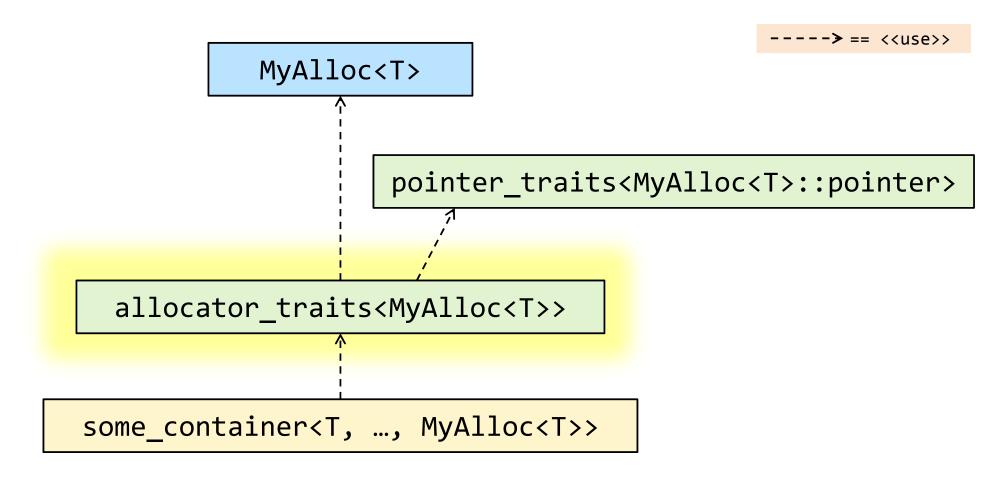
Modern C++ Allocator Model – Containers' Perspective



Example Container – Part 1

```
template<class T, class Alloc = allocator<T>>
class some container
  public:
   using value type
                    = T;
   using allocator type = Alloc;
   using alloc traits = std::allocator traits<allocator type>;
   using pointer
                   = typename alloc traits::pointer;
   using const pointer = typename alloc traits::const pointer;
   using size type
                        = typename alloc traits::size type;
   using difference type = typename alloc traits::difference type;
   using reference = value type&;
   using const reference = value_type const&;
   using iterator = implementation defined stuff;
   using const iterator = const implementation defined stuff;
};
```

Modern C++ Allocator Model – Allocator Traits



Allocator Traits – The Star of the Show

std::allocator_traits<Alloc>

- Provides a uniform interface to allocators used by the containers
 - Attempts to reflect certain important properties from a target allocator
- Picks sensible defaults for the properties and functions not defined by the target allocator
 - Supplies the boilerplate that was part of C++03 allocators
 - Backwardly-compatible with C++03 allocators
- Allows customization by using non-default properties and functions defined by the target allocator

```
template<class Alloc> struct allocator traits
   using allocator type
                        = Alloc:
                    = typename Alloc::value_type;
= INFERRED;
   using value_type
   using pointer
   using const pointer = INFERRED;
   using void pointer
                       = INFERRED;
   using const_void_pointer = INFERRED;
   using difference_type = INFERRED;
   using size type = INFERRED;
   using propagate on container copy assignment = INFERRED; //- POCCA
   using propagate on container move assignment = INFERRED; //- POCMA
                                              = INFERRED; //- POCS
   using propagate on container swap
                                              = INFERRED; //- IZEQ
   using is always equal
   template <class T> using rebind alloc = INFERRED;
   template <class T> using rebind traits = allocator traits<rebind alloc<T>>;
};
```

```
template<class Alloc> struct allocator traits
                     allocate(Alloc& a, size type n);
   static pointer
   static pointer
                     allocate(Alloc& a, size type n, const void pointer hint);
   static void
                     deallocate(Alloc& a, pointer p, size type n);
   template <class T, class... Args>
   static void
                     construct(Alloc& a, T* p, Args&&... args);
   template <class T>
    static void destroy(Alloc& a, T* p);
   static size type max size(Alloc const& a) noexcept;
   static Alloc
                     select on container copy construction(Alloc const& a);
};
```

```
template<class Alloc> struct allocator traits
    using pointer
                             = typename Alloc::pointer
                             | value type*;
    using const pointer
                             = Alloc::const pointer
                             | pointer traits<pointer>::rebind<const value type>;
    using void pointer
                             = Alloc::void pointer
                             pointer traits<pointer>::rebind<void>;
    using const void pointer = Alloc::const void pointer
                             pointer traits<pointer>::rebind<const void>;
    using difference type
                             = Alloc::difference type
                             | pointer traits<pointer>::difference type;
    using size type
                             = Alloc::size type
                             make unsigned t<difference type>;
```

```
template<class Alloc> struct allocator traits
    using propagate on container copy assignment
                                    = Alloc::propagate_on_container_copy_assignment
                                     | std::false type;
    using propagate on container move assignment
                                    = Alloc::propagate on container move assignment
                                     | std::false type;
    using propagate on container swap
                                    = Alloc::propagate on container swap
                                     | std::false type;
    using is always equal
                                    = Alloc::is always equal
                                     | std::is empty<Alloc>::type;
};
```

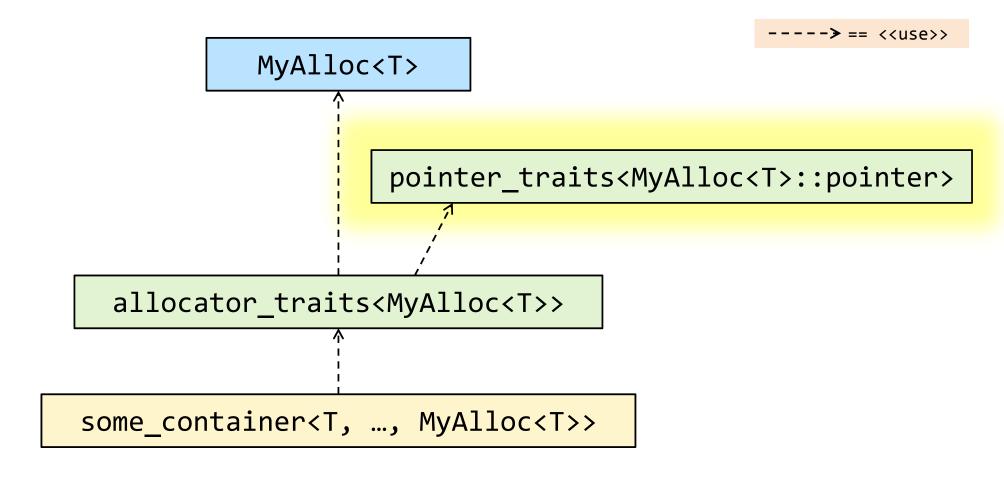
```
template<class Alloc> struct allocator traits
    template<class T> using rebind_alloc
                                            = Alloc::rebind<T>::other
                                             SomeAlloc<T,Args..> if Alloc is a SomeAlloc ill-formed;
    template<class T> using rebind_traits = allocator_traits<rebind_alloc<T>>;
```

```
template<class Alloc> struct allocator_traits
    static pointer
                      allocate(Alloc& a, size_type n)
                          return a.allocate(n);
    static pointer
                      allocate(Alloc& a, size_type n, const_void_pointer hint)
                          return a.allocate(n, hint);
                               | a.allocate(n);
                      deallocate(Alloc& a, pointer p, size_type n)
    static void
                          a.deallocate(p, n);
};
```

```
template<class Alloc> struct allocator traits
    template <class T, class... Args>
    static void construct(Alloc& a, T* p, Args&&... args)
                      a.construct(p, std::forward<Args>(args)...);
                       ::new (static_cast<void*>(p)) T(std::forward<Args>(args)...);
    template <class T>
    static void destroy(Alloc& a, T* p);
                      a.destroy(p);
                      p \rightarrow T();
};
```

```
template<class Alloc> struct allocator traits
    static size_type max_size(Alloc const& a) noexcept
                          return a.max size();
                               | numeric_limits<size_type>::max()/sizeof(value_type);
    static Alloc
                      select_on_container_copy_construction(Alloc const& a)
                          return a.select on container copy construction();
                               a;
};
```

Modern C++ Allocator Model – Pointer Traits



Pointer Traits

- Describes the properties of pointers and pointer-like types
- Provides the type of pointer differences
- Provide the pointed-to type
 - Given the type T*, get the type T
 - Given the type fancy_ptr<T>, get the type T
- Provide a transformation from one pointer type to another pointer type
 - Given the type T*, get the type U*
 - Given the type fancy_ptr<T>, get the type fancy_ptr<U>

Pointer-Like Types (AKA Fancy / Synthetic Pointers)

- Mentioned 4 times in the standard, but the only substance is in the requirements for NullablePointer (see Table 28 in N4687):
 - Must satisfy several requirements:
 - EqualityComparable, DefaultConstructible, CopyConstructible, CopyAssignable, and Destructible
 - Have swappable Ivalues
 - Default initialization my produce an indeterminate result; using may lead to UB
 - Value initialization produces a null result
 - Construction with / assignment from nullptr produces a null result
 - May be contextually converted to bool
 - Certain fundamental operations (see Table 28) may not throw exceptions

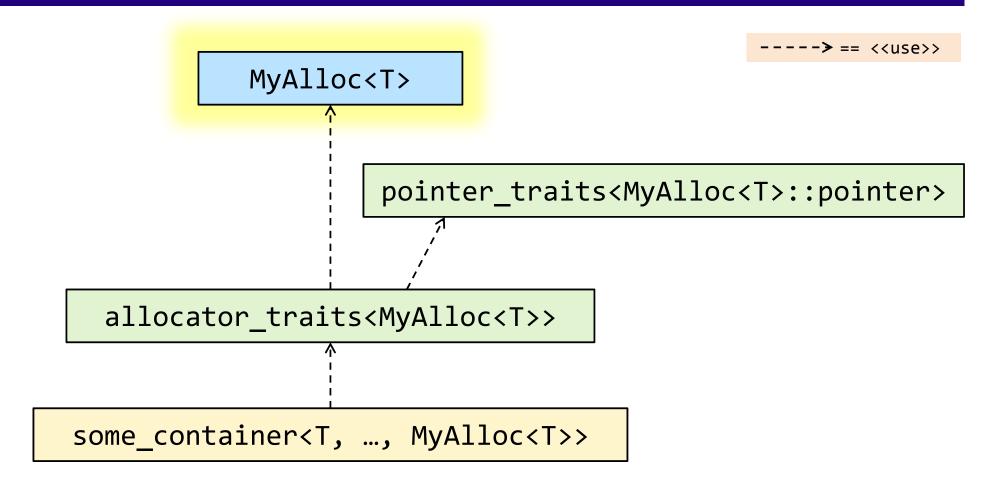
Pointer Traits

```
template<class T> struct pointer traits;
template<class T>
struct pointer traits<T*>
   using pointer = T*;
   using element_type = T;
   using difference_type = ptrdiff_t;
   template<class U> using rebind = U*;
   static pointer pointer_to(element_type& r) noexcept { return std::addressof(r); }
};
```

Pointer Traits

```
template<class Ptr>
struct pointer traits<Ptr>
    using pointer
                          = Ptr;
    using element type
                          = typename Ptr::element type
                            T if Ptr is a SomePointer<T,Args...>
                            ill-formed;
    using difference_type = typename Ptr::difference_type
                           | ptrdiff t;
    template<class U>
             using rebind = Ptr::template rebind<U>
                            SomePointer<U, Args...> if Ptr is a SomePointer<T, Args...>;
                            ill-formed
    static pointer pointer_to(element_type& r) { return Ptr::pointer_to(r); }
};
```

Modern C++ Allocator Model – The Allocator



A Minimal Allocator

```
template<class T>
struct minimal allocator
    using value type = T;
    minimal allocator(some params) {}
    template<class U> minimal allocator(minimal allocator<U> const&) noexcept {}
    T*
          allocate(size t n);
    void deallocate(T* p, size t);
};
template<class T> bool
operator ==(minimal allocator<T> const&, minimal allocator<T> const&);
template<class T> bool
operator !=(minimal allocator<T> const&, minimal allocator<T> const&);
```

The Modern Default Allocator

```
template<class T> class allocator
  public:
    using value type
                                                 = T;
    using propagate on container move assignment = true type;
    using is always equal
                                                 = true type;
    allocator() noexcept;
    allocator(const allocator&) noexcept;
    template<class U> allocator(const allocator<U>&) noexcept;
    ~allocator();
         allocate(size t n);
    T*
    void deallocate(T* p, size t n);
};
template<class T, class U>
bool operator ==(allocator<T> const&, allocator<U> const&) noexcept;
template<class T, class U>
bool operator !=(allocator<T> const&, allocator<U> const&) noexcept;
```

The Modern Default Allocator – Allocate / Deallocate

```
template<class T> inline T*
allocator<T>::allocate(size_t n)
    return static cast<T*>(::operator new(n * sizeof(T)));
template<class T> inline void
allocator<T>::deallocate(T* p, size_t)
    ::operator delete(p);
```

The Modern Default Allocator – Comparison

```
template<class T, class U> inline bool
operator ==(allocator<T> const&, allocator<U> const&) noexcept
    return true;
template<class T, class U> inline bool
operator !=(allocator<T> const&, allocator<U> const&) noexcept
    return false;
```

CppCon 2017 5-

There's a New Kid in Town

- Polymorphic memory resources (PMR)
 - Defined in namespace std::pmr
 - Provide runtime polymorphism with single type argument to containers
 - Client allocators store a pointer to a base class memory resource
 - No lateral propagation an allocator sticks for life

Polymorphic Memory Resource (PMR)

```
class pmr::memory resource
    static constexpr size t max align = alignof(max align t);
  public:
    virtual ~memory resource();
    void*
            allocate(size t bytes, size t alignment = max align);
    void
            deallocate(void* p, size t bytes, size t alignment = max align);
    bool
            is equal(const memory resource& other) const noexcept;
  private:
    virtual void*
                    do allocate(size t bytes, size t alignment) = 0;
                    do deallocate(void* p, size t bytes, size t alignment) = 0;
    virtual void
                    do is equal(const memory resource& other) const noexcept = 0;
    virtual bool
};
```

Polymorphic Memory Resource (PMR)

```
class monotonic buffer resource: public memory resource
  public:
    monotonic buffer resource();
    explicit monotonic buffer resource(memory resource *upstream);
    monotonic buffer resource(const monotonic buffer resource&) = delete;
    virtual ~monotonic buffer resource();
    monotonic buffer resource& operator=(const monotonic buffer resource&) = delete;
    void
                        release();
    memory resource*
                        upstream resource() const;
  protected:
            do allocate(size t bytes, size t alignment) override;
    void*
            do deallocate(void* p, size t bytes, size t alignment) override;
    void
            do is equal(const memory resource& other) const noexcept override;
    bool
};
```

Polymorphic Allocator

```
template<class T> class polymorphic allocator
    memory resource*
                        memory rsrc;
  public:
    using value type = T;
    polymorphic allocator() noexcept;
    polymorphic allocator(memory resource* r);
    polymorphic allocator(const polymorphic allocator& other) = default;
    template <class U>
    polymorphic allocator(const polymorphic allocator<U>& other) noexcept;
    polymorphic allocator& operator=(const polymorphic allocator& rhs) = delete;
    T*
            allocate(size t n);
            deallocate(T* p, size t n);
    void
                           select on container copy construction() const;
    polymorphic allocator
};
```

PMR Type Aliases

```
//- A partial list of type aliases now provided in std::pmr
namespace pmr
    template <class charT, class traits = char traits<charT>>
    using basic string = std::basic string<charT, traits, polymorphic allocator<charT>>;
    using string = basic string<char>;
    template <class T>
    using deque = std::deque<T, polymorphic allocator<T>>;
    . . .
    template <class Key, class Compare = less<Key>>
    using set = std::set<Key, Compare, polymorphic allocator<Key>>;
    template <class Key, class Hash = hash<Key>, class Pred = equal to<Key>>
    using unordered set = std::unordered set<Key, Hash, Pred,</pre>
                                              polymorphic allocator<Key>>;
```

A Container's Point of View

Example Container – Part 2

```
template<class T, class Alloc = allocator<T>>
class some container
  public:
   using value type
                    = T;
   using allocator type = Alloc;
   using alloc traits = std::allocator traits<allocator type>;
   using pointer
                   = typename alloc traits::pointer;
   using const pointer = typename alloc traits::const pointer;
   using size type = typename alloc traits::size type;
   using difference type = typename alloc traits::difference type;
   using reference = value type&;
   using const reference = value_type const&;
   using iterator = implementation defined stuff;
   using const iterator = const implementation defined stuff;
};
```

Example Container – Part 2

```
template<class T, class Alloc<T>>
class some container
  public:
    some container(some container const& other);
    some container(some container const& other, allocator type const& alloc);
    some container(some container&& other);
    some_container(some_container&& other, allocator_type const& alloc);
    some container& operator =(some container const& other);
    some container& operator =(some container&& other);
    swap(some container& other);
};
```

Example Container – Part 2

```
template<class T, class Alloc<T>>
class some_container
  private:
    impl_data
                   m_impl;
    allocator_type m_alloc;
    template<class Iter>
            assign_from(Iter f, Iter 1);
    void
            clear_and_deallocate_memory()
    void
};
```

Example Container – Copy Construction

```
template<class T, class Alloc>
some container<T,Alloc>::some container(some container const& other)
   m impl()
    m alloc(traits type::select on container copy construction(other.m alloc))
    this->assign from(other.cbegin(), other.cend());
template<class T, class Alloc>
some container<T, Alloc>::some container(some container const& other,
                                        allocator type const& alloc)
   m impl()
    m alloc(alloc)
    this->assign from(other.cbegin(), other.cend());
```

CppCon 2017 6-

Example Container – Move Construction

```
template<class T, class Alloc>
some container<T,Alloc>::some container(some container&& other)
    m impl(std::move(other.m impl)
    m alloc(std::move(other.m alloc))
{}
template<class T, class Alloc>
some container<T,Alloc>::some container(some container&&
                                                               other,
                                         allocator type const& alloc)
   m_impl(std::move(other.m_impl)
    m alloc(alloc)
{}
```

Example Container – Copy Assignment

```
template<class T, class Alloc> some container&
some container<T,Alloc>::operator =(some container const& other)
    if (&other != this)
        if (alloc traits::POCCA == std::true type)
            if (this->m_alloc != other.m_alloc)
                this->clear and deallocate memory();
            this->m alloc = other.m alloc;
        this->assign from(other.cbegin(), other.cend());
    return *this;
```

Example Container – Move Assignment

```
template<class T, class Alloc> some container&
some container<T,Alloc>::operator =(some container&& rhs)
    if (alloc traits::POCMA == std::true type)
        this->clear and deallocate memory();
        this->m alloc = std::move(rhs.m alloc);
        this->m impl = std::move(rhs.m impl);
   else if (alloc traits::is always equal == std::true type
            this->m alloc == rhs.m alloc)
       this->clear and deallocate memory();
        this->m impl = std::move(rhs.m impl);
   else
       this->assign(std::move iterator(rhs.begin()), std::move iterator(rhs.end()));
   return *this;
```

Example Container – Swap

```
template<class T, class Alloc> void
some container<T,Alloc>::swap(some container& other)
    if (&other != this)
        if (alloc traits::POCS == std::true type)
            std::swap(this->m impl, other.m impl);
            std::swap(this->m alloc, other.m alloc);
        else if (alloc traits::is always equal == std::true type
                 this->m alloc == rhs.m alloc)
            std::swap(this->m impl, other.m impl);
        else
               //- POCS is std::false type and this->m alloc != other.m alloc
                //- This is undefined behavior.
```

A Few Guidelines

Thinking About Memory Allocation – Design Decisions

- Plumbing
 - Public Interface: Traditional or PMR
- Structural Management
 - Addressing Model
 - Storage Model
 - Pointer Interface
 - Allocation Strategy
- Concurrency Management
 - Thread Safety
 - Transaction Safety

Thinking About Your New Allocator – Traditional

- Review the minimal and default allocator interfaces
- Decide how to specify properties and functionality
 - Everything in the allocator type; or,
 - Partially specializing std traits types
- Synthetic pointers or ordinary pointers?
 - If synthetic, don't forget your pointer type's re-binder
 - If synthetic, you must define a nested to_pointer() member function
- Do specify your allocator's nested re-binder
- Do specify your allocator's nested value_type alias

Thinking About Your New Allocator – Traditional

- If instances always compare equal?
 - Define a nested type alias is_always_equal set to true_type
- If instances do not propagate on copy construction
 - Define select_on_container_copy_construction() that returns a temporary value of your allocator type
- If instances propagate on copy assignment,
 - Define nested alias propagate_on_container_copy_assignment as true_type
- If instances propagate on move assignment,
 - Define nested alias propagate_on_container_move_assignment as true_type
- If instances propagate on swap,
 - Define nested alias propagate_on_container_swap as true_type

Thinking About Your New Allocator - PMR

- Which pmr base class?
 - pmr::memory_resource
 - There are others, derived from pmr::memory_resource ...
 - synchronized_pool_resource
 - unsynchronized_pool_resource
 - monotonic_buffer_resource

Thinking About Your New Allocator

- If you don't need fancy pointers, and you can live with a little extra overhead in each container
 - Create a new PMR type, derived from pmr::memory_resource
- Otherwise, you'll need to jump into the deep end of the traditional allocator pool

Thank You for Attending!

https://gitlab.com/BobSteagall/talks/CppCon2017 (that's GitLab)