# An intro to Boost.PolyCollection



using std::cpp 2017

Joaquín M López Muñoz <joaquin.lopezmunoz@gmail.com> Madrid, November 2017

# Before the show begins: some news from the Boost front



# Before the show begins: some news from the Boost front

- 145 libraries (1.65.1) and counting, 5 new last year
  - Fiber, QVM, Process, PolyCollection, Stacktrace
- Trends
  - C++11 or later increasingly required by new libs
  - Postmodern (meta)functional: Hana, Fit, Mp11
  - OS and communications: Beast, Fiber, Process, Stacktrace
- Building Boost
  - Boost.Build challenged by CMake
  - Bincrafters now provides Boost packages through conan
- You have no excuses for not using Boost!

# Let 's dive into Boost.PolyCollection



# Type hierarchy for a role playing game

```
struct sprite
{
   sprite(int id):id(id){}
   virtual ~sprite()=default;
   virtual void render(std::ostream& os)const=0;
   int id;
};
```

```
struct warrior:sprite
{
  using sprite::sprite;
  warrior(std::string rank,int id):
    sprite{id},rank{std::move(rank)}{}
  void render(std::ostream& os)const override
  {os<<rank<<" "<<id;}
  std::string rank="warrior";
};</pre>
```

```
struct juggernaut:warrior
{
  juggernaut(int id):
    warrior{"juggernaut",id}{}
};
```

```
struct goblin:sprite
{
  using sprite::sprite;
  void render(std::ostream& os)
  const override
  {os<<"goblin "<<id;}
};</pre>
```

## Populate and render

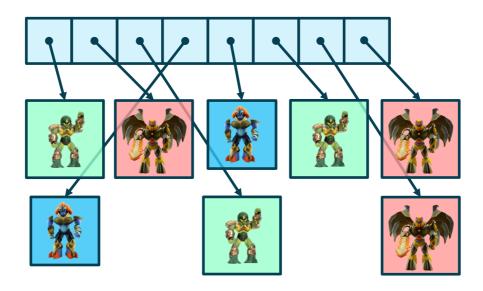


```
std::vector<std::unique ptr<sprite>> c;
                          gen{92748}; // some arbitrary random seed
std::mt19937
std::discrete distribution<> rnd{{1,1,1}};
switch(rnd(gen)){
   // watch out: std::make unique requires C++14
   case 0: c.push back(std::make unique<warrior>(i));break;
   case 1: c.push_back(std::make_unique<juggernaut>(i));break;
   case 2: c.push back(std::make unique<goblin>(i));break;
const char* comma="";
for(const auto& p:c){
 std::cout<<comma;</pre>
 p->render(std::cout);
 comma=",";
std::cout<<"\n";</pre>
```

# Piece of cake



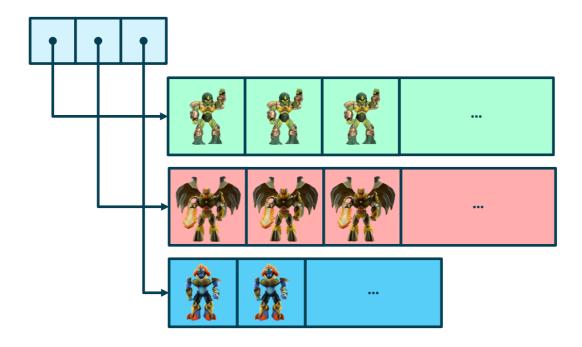
# There are hidden inefficiencies, though



- Objects are scattered through memory -> CPU cache misses
- Types are mixed up along traversal → CPU branch prediction fails
  - We could sort the vector on type\*, object scattering is harder to fix though
- All boils down to OOP requiring pointer indirections

\*See my "Mind the cache" talk at github.com/joaquintides/usingstdcpp2015

# Packing things together



- boost::base\_collection stores elements in per-type dedicated segments
- [begin(), end()) traverses each segment in succession
  - So, free ordering of elements is lost (except within segment)
- In return, performance increases spectacularly

```
boost::base collection<sprite> c;
                              gen{92748}; // some arbitrary random seed
std::mt19937
std::discrete distribution<> rnd{{1,1,1}};
for(int i=0;i<8;++i){</pre>
                             // assign each type with 1/3 probability
  switch(rnd(gen)){
    case 0: c.insert(warrior{i});break;
    case 1: c.insert(juggernaut{i});break;
    case 2: c.insert(goblin{i});break;
const char* comma="";
for(const sprite& s:c){
  std::cout<<comma;</pre>
  s.render(std::cout);
  comma=",";
std::cout<<"\n";</pre>
```

■ Spot the differences with the std::vector code

Insertion interface mimics std::multi\_set rather than std::vector

**Emplacement requires that type of element be specified** 

```
std::sort( // sort warriors in descending id order
   c.begin
c.begin
c.begin
warrior>
(), c.end
warrior>
(),

[](const warrior& x,const warrior& y){return x.id>y.id;});

for(const warrior& w:c.segment
warrior>
()) // print warrior ids
   std::cout<<w.id<<"\n";

c.erase(c.begin(typeid(warrior))); // erase first warrior

c.clear</pre>
c.clear
c.clear
warrior>
(); // erase all warriors
```

- Watch out: begin(typeid(warrior) is not the same as begin<warrior>()
- Two kinds of local iterators (prefer the latter)

Insertion interface mimics std::multi\_set rather than std::vector

**Emplacement requires that type of element be specified** 

Emplacement requires that type of element be specified

Insertion interface mimics std::multi\_set rather than std::vector

### Emplacement requires that type of element be specified

```
c.emplace<goblin>(11);
```

c.emplace\_pos<goblin>(c.begin<goblin>(),12); // 1st of goblins

Insertion interface mimics std::multi\_set rather than std::vector

Emplacement requires that type of element be specified

```
// new assortment of sprites
std::array<std::unique_ptr<sprite>,3> a{{
    std::make_unique<elf>(13),
    std::make_unique<ghoul>(14),
    std::make_unique<amazon>(15)
}};
c.register_types<elf,ghoul,amazon>(); // otherwise exception is thrown below
for(const auto&p:a)c.insert(*p);
```

Insertion interface mimics std::multi\_set rather than std::vector

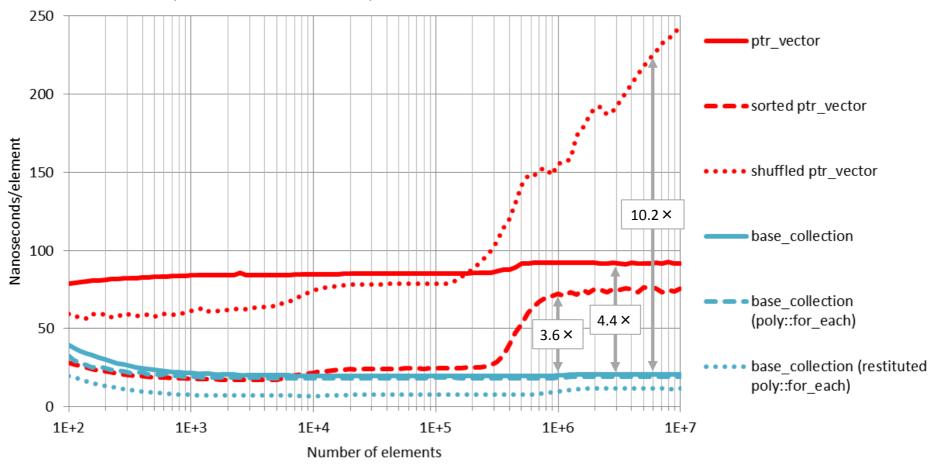
**Emplacement requires that type of element be specified** 

**New segments are created automagically; otherwise, use register\_types** 

All of this is well, but is it really faster?

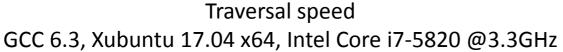
### Performance

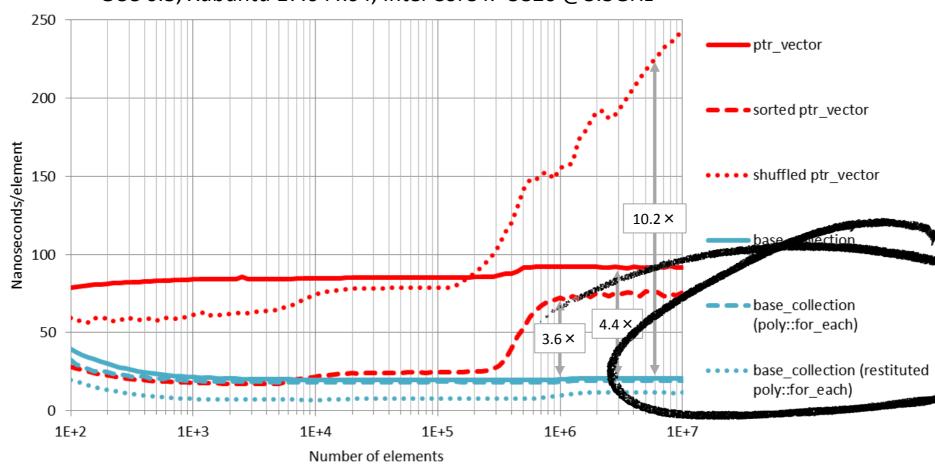
Traversal speed GCC 6.3, Xubuntu 17.04 x64, Intel Core i7-5820 @3.3GHz



Full results at boost.org/doc/html/poly\_collection/performance.html

### Performance





Full results at boost.org/doc/html/poly\_collection/performance.html Which is fastest?

```
for(const sprite& s:c){
   std::cout<<comma;
   s.render(std::cout);
   comma=",";
}

for(const auto& seg_info:c.segment_traversal()){
   for(const sprite& s:seg_info){
      std::cout<<comma;
      s.render(std::cout);
      comma=",";
   }
}</pre>
```

 Double loop the basis of Boost.PolyCollection dedicated algorithms

```
boost::poly_collection::for_each(c.begin(),c.end(),[&](const sprite& s){
   std::cout<<comma;
   s.render(std::cout);
   comma=",";
});</pre>
```

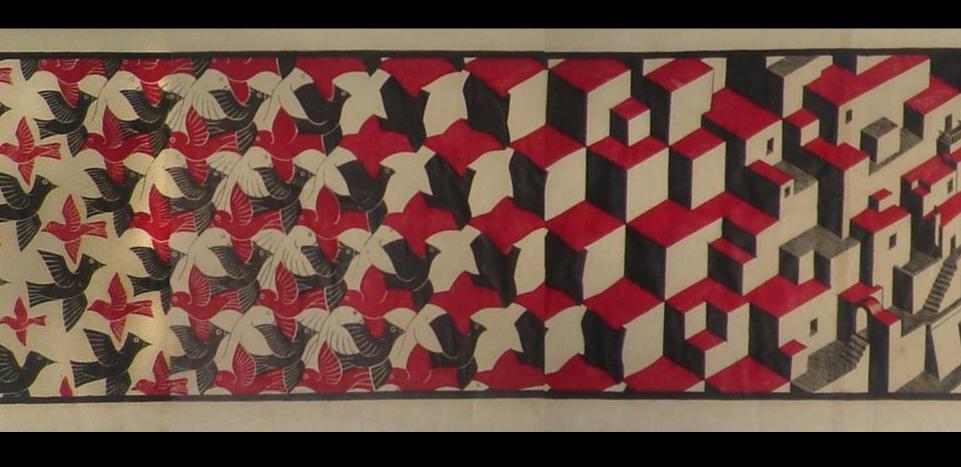
■ 31 algorithms adapted from <algorithm>

```
boost::poly_collection::for_each<warrior,juggernaut,goblin>(
   c.begin(),c.end(),[&](const auto& s){
      std::cout<<comma;
      s.render(std::cout);
      comma=",";
   }
);</pre>
```

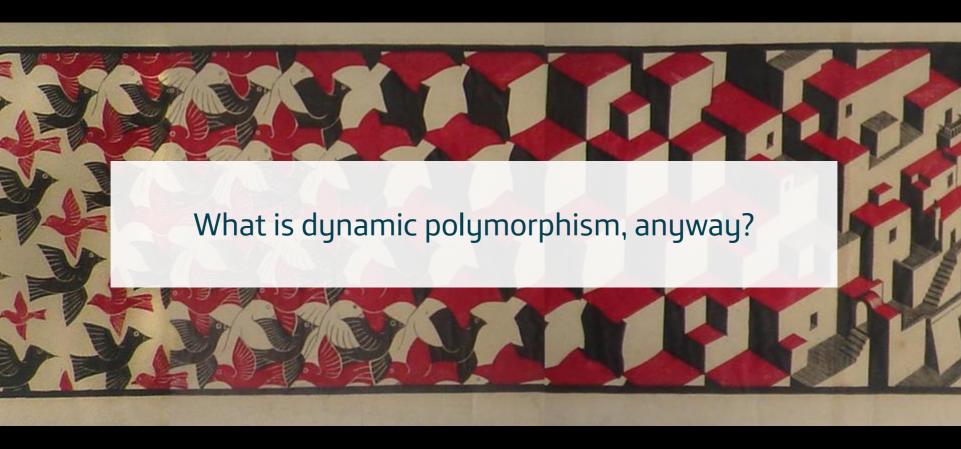
- Note the generic lambda
- **Type restitution** is the inverse of type abstraction

```
warrior* pw=new warrior{16};
sprite* ps=pw; // abstraction
warrior* pw2=static_cast<warrior*>(ps); // restitution
```

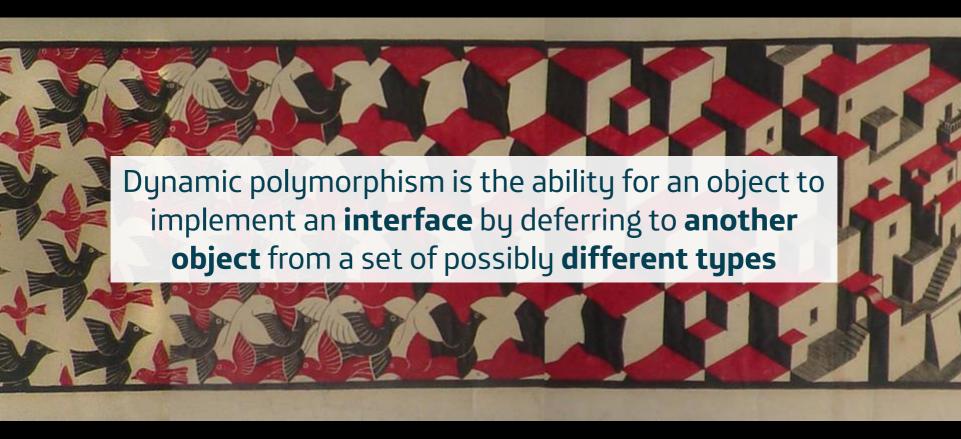
# A new look on dynamic polymorphism



# A new look on dynamic polymorphism



# A new look on dynamic polymorphism



# OOP is *just one model* of dynamic polymorphism...

```
sprite(int id);
virtual ~sprite();
virtual void render(std::ostream& os)const;
                                                 interface
int id;
                                               superobject
warrior(std::string rank,int id);
void render(std::ostream& os)const override;
std::string rank;
                                                subobject
```

# ...and not a particularly flexible one at that

- Implementation and subtyping are coupled together by inheritance
- Implementation is intentional (again, through inheritance)
- Intersecting interfaces require forethought and virtual inheritance
  - dynamic\_cast comes into play
  - Ever heard of the diamond problem?\*
  - Believe me, you just don't want to enter this hell

# Type erasure is the new cool...

...and std::function its most accessible example

```
int foo(int x){return x;}
struct bar{int operator()(int x)const{return 2*x;}};
using signature=int(int);
std::function<signature> f=&foo,g=bar{},h=[](int x){return 3*x;};
std::cout<<f(1)+g(2)+h(3)<<"\n";</pre>
```

This is a new model of dynamic polymorphism

■ Interface: signature

■ Superobject: std::function<signature>

■ Subobjects: int(\*)(int), bar, compatible lambdas

Interface compliance is checked at compile time

```
boost::function collection<int(int)> c;
                        gen{92748}; // some arbitrary random seed
std::mt19937
std::discrete distribution<> rnd{{1,1,1}};
switch(rnd(gen)){
   case 0: c.insert(&foo);break;
   case 1: c.insert(bar{});break;
   case 2: c.insert([](int x){return 3*x;});break;
int res=0;
for(const auto& f:c)res+=f(1);
std::cout<<res<<"\n";</pre>
```

- Same interface as boost::base\_collection
- How's this different from std::vector<std::function<int(int)>>?

# Duck typing: type erasure on steroids

```
std::ostream& operator<<(std::ostream& os,const sprite& s)
{s.render(os);return os;}

using concept_=boost::mpl::vector<
   boost::type_erasure::ostreamable<>,
   boost::type_erasure::copy_constructible<>
>;
boost::type_erasure::any<concept_> a=5,b=std::string{"hello"},c=warrior{16};

std::cout<<a<<", "<<b<<", "<<c<<"\n";</pre>
```

- Interface: concept\_
- Superobject: boost::type\_erasure::any<concept\_>
- Subobjects: anything satisfying concept\_
- Boost.TypeErasure not the only duck typing framework: see for instance Louis Dionne's **Dyno**
- Ultimate solution: native run-time C++ concepts (some day)

```
boost::any collection<boost::type erasure::ostreamable<>> c;
                          gen{92748}; // some arbitrary random seed
std::mt19937
std::discrete distribution<> rnd{{1,1,1,1,1}};
switch(rnd(gen)){
   case 0: c.insert(warrior{i});break;
   case 1: c.insert(juggernaut{i});break;
   case 2: c.insert(goblin{i});break;
   case 3: c.insert(boost::format{"message %1%"}%i);break;
   case 4: c.insert(i);break;
const char* comma="";
for(const auto& x:c){
 std::cout<<comma<<x;</pre>
 comma=",";
std::cout<<"\n";</pre>
```

■ Same interface as before... you get the idea

# So many butterflies, so little time

# So many butterflies, so little time

- Boost.PolyCollection packs elements of the same type together for maximum performance
  - Classic OOP boost::base\_collection
  - Callable entities boost::function\_collection
  - Duck typing boost::any\_collection
- Remember: order of elements no longer free
- Squeeze more speed: dedicated algorithms and type restitution
- Dynamic polymorphism goes well beyond OOP
  - Read, play, use
- Drop me a note when you use Boost.PolyCollection

# An intro to Boost.PolyCollection

# Thank you

boost.org/libs/poly\_collection github.com/joaquintides/usingstdcpp2017

using std::cpp 2017

Joaquín M López Muñoz <joaquin.lopezmunoz@gmail.com> Madrid, November 2017