

Расширения стандартной  
библиотеки Си++ для  
многопоточного  
программирования в стандарте  
Си++ 11



# Новый стандарт Си++

19 декабря 2011 года международная организация по стандартизации ISO объявила о выходе стандарта C — ISO/IEC 9899:2011 (ранее неофициально известного как C1X).

## Основные изменения по сравнению с C98 и C03:

Тип auto

Диапазонный for

Управление поведением по умолчанию: default и delete

Rvalue ссылки

Лямбда-выражения

**Поддержка многопоточности**

**future & promise**

...



# Поддержка стандарта транслятором GNU C++

Версия GNU C++ 4.7.2 практически полностью поддерживает параллельные возможности Си++ 11.

Сборка приложения

```
g++ -std=c++11 -o ex2 ex2.cpp -pthread
```

В более ранних версиях

```
g++ -std=c++0x
```



# Поддержка многопоточности

Класс **std::thread**,

заголовочный файл **<thread>**



# Простейшая параллельная программа

```
#include <iostream>
#include <thread>

void hello()
{
    std::cout<<"Hello Concurrent World\n";
}

int main()
{
    std::thread t(hello);
    t.join();
}
```



# Запуск потока с экземпляром класса

```
#include <iostream>
#include <thread>

int main()
{
    class background_task
    {
    public:
        void operator()()
        {
            std::cout << "Hello\n";
        }
    };
```



# Запуск потока с экземпляром класса

```
background_task f;  
    std::thread my_thread(f);  
    my_thread.join();  
}
```

В новом потоке создается копия экземпляра класса!



# Функциональные объекты в C++ 11

- Начиная со стандарта C++11 шаблонный класс `std::function` является полиморфной обёрткой функций для общего использования.
- Объекты класса `std::function` могут хранить, копировать и вызывать произвольные вызываемые объекты - функции, лямбда-выражения, выражения связывания и другие функциональные объекты.
- В любом месте, где необходимо использовать указатель на функцию для её отложенного вызова, или для создания функции обратного вызова, вместо него может быть использован `std::function`, который предоставляет пользователю большую гибкость в реализации.





# Функциональные объекты

```
#include <stdio.h>
```

```
#include <functional>
```

```
int sum(int a, int b) {  
    return a + b;  
}
```

```
main() {  
    std::function<int (int, int)> f = sum;  
    printf("1 + 2 = %d\n", f(1, 2));  
}
```



# std::bind

функция `std::bind` называется “привязывает” некоторые аргументы к функциональному объекту, создавая новый функциональный объект. То есть, вызов `std::bind` эквивалентен вызову функционального объекта с некоторыми определёнными параметрами.

Передавать можно или непосредственно значения аргументов, или специальные имена, определенные в пространстве имен `std::placeholders`.



# std::bind

```
template< class F, class... Args >  
/*unspecified*/ bind( F&& f, Args&&... args );
```

Return value:

A function object of unspecified type T, for which

**std::is\_bind\_expression<T>::value == true,**  
( объект получен с помощью std::bind)

and which can be stored in std::function



# Пример std::bind

```
#include <stdio.h>
#include <functional>
```

```
int sum(int a, int b) {
    return a + b;
}
```

```
main() {
    std::function<int (int, int)> f = sum;
    printf("1 + 2 = %d\n", f(1, 2));
    auto g = std::bind(sum, 1, std::placeholders::_1);
    printf("1 + 4 = %d\n", g(4));
}
```

Специальный объект,  
выделяющий параметр для  
последующего «связывания»

Неспецифицируемый тип,  
вычисляемый в процессе  
компиляции



# Копирование параметров при передаче в качестве аргументов функции

Параметры всегда копируются и сохраняются во временных объектах перед передачей при использовании **std::bind**. Чтобы этого не происходило, следует использовать конструкцию **std::ref**, позволяющую сделать из объекта ссылку на него, которую можно копировать.



# Передача объекта класса по ссылке

Чтобы предотвратить копирование объекта, следует использовать тип **`std::reference_wrapper<T>`**, который является «оберткой» вокруг ссылки типа `T` и возвращает ссылку на объект посредством метода `get()` либо в результате неявного преобразования типа (может использоваться в контексте, где требуется обычная ссылка)

Создается с помощью **`std::ref()`** и **`std::cref()`**.

Этот класс обычно используется для явной передачи объектов ссылочного типа:

```
template <typename T> void foo (T val);  
int x;
```

```
foo (std::ref(x)); // x – ссылка типа int&
```

```
foo (std::cref(x)); // x – ссылка типа const int &
```



# Пример использования reference\_wrapper

```
#include <iostream>
#include <functional>
```

```
class A {
public:
    void f() {
        std::cout << "Hello, A\n";
    }
};
```

```
void process(A& a) {
    a.f();
}
```

```
int main() {
    A a;
    std::reference_wrapper<A> ra = std::ref(a);
    process(ra);
    return 0;
}
```

```
posypkin@mikhail:~/exp/sandbox/c11$ g++ -std=c++11 -o ref ref.cpp
posypkin@mikhail:~/exp/sandbox/c11$ ./ref
Hello, A
posypkin@mikhail:~/exp/sandbox/c11$
```

# Пример использования reference\_wrapper

```
#include <iostream>
#include <vector>

class A {
public:

    void f() {
        std::cout << "Hello, A\n";
    }
};

void process(A& a) {
    a.f();
}
```

```
int main() {
    A a;
    std::reference_wrapper<A> ra = std::ref(a);
    std::vector< std::reference_wrapper<A> >
    vra;
    vra.push_back(ra);
    vra.push_back(ra);
    vra.push_back(ra);
    for(auto o : vra) {
        process(o);
    }
    return 0;
}
```

```
posypkin@mikhail:~/exp/sandbox/c11$ c++ -std=c++11 -o vref vref.cpp
posypkin@mikhail:~/exp/sandbox/c11$ ./vref
Hello, A
Hello, A
Hello, A
posypkin@mikhail:~/exp/sandbox/c11$
```



# Передача аргументов по ссылке и по значению

```
#include <stdio.h>
#include <functional>
```

```
class C {
public:
    C(int v) : a(v) {}
    int a;
};
```

```
main() {
    C c1(1);
    C& rc1 = c1;
    C c2(2);
    C& rc2 = c2;
```

```
#if 1
    std::reference_wrapper<C> rc3 = c1;
    rc3 = std::ref(rc2);
    rc3.get().a = 10;
    printf("c1.a = %d\n", c1.a);
    printf("c2.a = %d\n", c2.a);
#else
    C& rc3 = c1;
    rc3 = rc2;
    rc3.a = 10;
    printf("c1.a = %d\n", c1.a);
    printf("c2.a = %d\n", c2.a);
#endif
}
```



# Использование обычных ссылок

```
C& rc3 = c1;  
rc3 = rc2;  
rc3.a = 10;  
printf("c1.a = %d\n", c1.a);  
printf("c2.a = %d\n", c2.a);
```

напечатает (присваивание ссылок приводит к присваиванию значений объектов)

```
c1.a = 10  
c2.a = 2
```



# Использование reference\_wrapper

```
std::reference_wrapper<C> rc3 = c1;  
rc3 = std::ref(rc2);  
rc3.get().a = 10;  
printf("c1.a = %d\n", c1.a);  
printf("c2.a = %d\n", c2.a);
```

напечатает (присваивание ссылок приводит к изменению  
указываемого объекта)

```
c1.a = 1  
c2.a = 10
```



# Пример использования ref для передачи объекта по ссылке

```
class F {  
public:  
    F() : a(10) {}  
  
    void operator()() {  
        a = 5;  
    }  
    int a;  
};
```



# Пример использования ref для передачи объекта по ссылке

```
main(int argc, char* argv[]) {  
    F f;  
    std::thread th(f);  
    th.join();  
    printf("a = %d\n", f.a);  
  
    std::thread thr(std::ref(f));  
    thr.join();  
    printf("a = %d\n", f.a);  
}
```

```
mikdebian@debian32: ~/exp/course/c11$ ./thref  
a = 10  
a = 5  
mikdebian@debian32: ~/exp/course/c11$
```



# Передача аргументов

Передача аргументов в создаваемый поток производится также просто, как и при обычном вызове функции, например:

```
void f(int i, std::string const& s);  
std::thread t(f, 3, "hello");
```

При этом необходимо помнить, что аргументы **копируются** в запускаемый поток. Игнорирование этого факта может служить источником ошибок.



# Возможные скрытые источники ошибок

```
void f(int n, std::string& s);
```

```
void oops(int some_param)
{
    char buffer[1024];
    sprintf(buffer,"%i",some_param);
    std::thread t(f,3,buffer);
    t.detach();
}
```

Преобразование типа производится в созданном потоке, при этом локальная переменная **buffer** может перестать существовать к этому моменту, что вызовет ошибку



# Корректный код

```
void not_oops(int some_param)
{
    char buffer[1024];
    sprintf(buffer,"%i",some_param);
    std::thread t(f,3,std::string(buffer));
    t.detach();
}
```

Явное преобразование  
производится до создания  
экземпляра объекта  
**std::thread** в потоке, в новый  
поток копируется полностью  
объект типа **std::string**





# Передача аргументов по ссылке

```
int a = 5;
main(int argc, char* argv[]) {
    std::thread th(f, a);
    th.join();
    printf("a = %d\n", a);

    std::thread thr(f, std::ref(a));
    thr.join();
    printf("a = %d\n", a);
}
```

```
mikdebian@debian32:~/exp/course/c11$ ./thrparref
a = 5
a = 10
mikdebian@debian32:~/exp/course/c11$ █
```



# Получение значения функции, работающей в потоке

- Результат, возвращаемый функциональным объектом-аргументом потока, игнорируется.
- Для передаче родительскому потоку возвращаемого значения, нужно использовать другие конструкции, например `std::future`

# Потоки- «демоны»

- Метод **join()** класса **std::thread** позволяет дождаться завершения созданного потока. В этом методе происходит освобождение ресурсов, занятых потоком.
- Метод **detach()** разрывает ассоциацию объекта типа **std::thread** и созданного потока, переменная может использоваться повторно для создания других потоков. Освобождение занятых ресурсов берет на себя библиотека Си ++ и выполняется после завершения потока.
- Все потоки приложения завершаются при завершении приложения.



# Пример создания потоков-«демонов»

```
class background_task
{
public:
    void operator()(char c)
    {
        for(int i = 0; i < 10; i++) {
            std::cout << c << std::flush;
            sleep(1);
        }
        std::cout << "thread [" << c << "] finished\n" << std::flush;
    }
};
```



# Пример создания потоков- «демонов»

```
main() {  
    background_task f;  
    std::cout << "Enter y to terminate\n" << std::flush;  
    while(true) {  
        char c;  
        std::cin >> c;  
        if(c == 'y')  
            break;  
        else {  
            std::thread my_thread(f, c);  
            my_thread.detach();  
        }  
    }  
}
```



# Результат работы программы

Enter y to terminate

a

aaaaba

babababcb

cab

thread [a] finished

bcbcbcb

thread [b] finished

cccc

thread [c] finished

y



# Копирование экземпляров класса `std::thread`

Для сокращения числа возможных ошибок не разрешается иметь более одной ссылки на объект типа **`std::thread`**. Поэтому **копирование** запрещено, разрешено только **перемещение** объекта.



# std::move

**std::move** – функция, преобразующая тип **lvalue** (T&) к типу **rvalue** (T&&). Это позволяет перегружать конструкторы по умолчанию и операторы присваивания таким образом, чтобы упростить перемещение объекта.

Применение **std::move** указывает на то, что аргумент не может далее использоваться и должен «умереть» после завершения присваивания. Чтобы избежать двойного удаления аллоцированные объекты конструктор перемещения или оператор присваивания с семантикой перемещения обнуляет аллоцируемые объекты перемещаемого класса (правой части присваивания) или переписывает туда указатели на объекты изменяемого класса (левая часть присваивания).

Обычно имеют два конструктора и оператора присваивания (для lvalue и rvalue).

C(C && c)        // перемещение

C(const C & c)    // копирование





# Неправильное использование rvalue

```
class A {  
    char* p;  
  
    ...  
    A(A && a) {  
        delete [] p;  
        p = a.p;  
        //последующий вызов  
        delete  
        //в деструкторе класса a  
        //приведет к ошибке  
    }
```



# Правильное использование **rvalue**

## 1. Обнуление

```
class A {  
    char* p;  
    ...  
    A(A && a) {  
        delete [] p;  
        p = a.p;  
        a.p = nullptr; //  
        //последующий вызов delete  
        //в деструкторе класса a не  
        //приведет к ошибке  
    }
```

## 2. Обмен указателями

```
class A {  
    char* p;  
    ...  
    A(A && a) {  
        char *tmp = p;  
        p = a.p;  
        a.p = tmp; // будет  
        // освобожден при  
        // вызове деструктора a  
    }
```



# Конструкторы и операторы присваивания `std::thread`

## `std::thread::thread`

---

<code>thread();</code>	(1)	(since C++11)
<code>thread( thread&amp;&amp; other );</code>	(2)	(since C++11)
<code>template&lt; class Function, class... Args &gt; explicit thread( Function&amp;&amp; f, Args&amp;&amp;... args );</code>	(3)	(since C++11)
<code>thread(const thread&amp;) = delete;</code>	(4)	(since C++11)

## `std::thread::operator=`

---

<code>thread&amp; operator=( thread&amp;&amp; other );</code>	(since C++11)
---	---------------

Assigns the state of `other` to `*this` using move semantics.

If `*this` still has an associated running thread (i.e. `joinable() == true`), `std::terminate()` is called



# Копирование экземпляров класса `std::thread`

```
std::thread my_thread(f);  
std::thread new_thread;
```

Неправильно:

```
new_thread = my_thread;
```

Правильно:

```
new_thread = std::move(my_thread);
```



# Возврат потока функцией – корректный вариант

```
std::thread make_thread() {  
    background_task f;  
    std::thread my_thread(f);  
    return my_thread;  
}  
main() {  
    std::thread new_thread;  
    new_thread = make_thread();  
    new_thread.join();  
}
```

Возвращается  
**rvalue**



# Запуск группы потоков

```
#include <thread>
#include <iostream>
#include <vector>

void bt() {
    printf("Hello\n");
}

main() {
    std::vector<std::thread> v;
    for(int i = 0; i < 8; i++) {
        std::thread th(bt);
        v.push_back(std::move(th));
    }
    for(auto& t : v) t.join();
}
```

```
admin@irbis8:~/exp/lectures$ ./ex5
Hello
Hello
Hello
Hello
Hello
Hello
Hello
Hello
```



# Идентификация потоков

Идентификатор имеет тип **`std::thread::id`** и является уникальным.

Методы для получения идентификатора:

Метод **`get_id()`** класса **`std::thread`** позволяет получить идентификатор потока, соответствующего заданному классу. Метод **`std::this_thread::get_id()`** возвращает экземпляр вызывающего потока.



# Идентификация потоков

```
std::thread::id master_id;
```

```
void bt() {  
    if(std::this_thread::get_id() == master_id)  
        printf("Hello from master thread\n");  
    else  
        printf("Hello\n");  
}
```





# Идентификация потоков

```
main() {  
    std::vector<std::thread> v;  
    int n = 8;  
    for(int i = 0; i < n; i++) {  
        std::thread th(bt);  
        if(i == 0) master_id = th.get_id();  
        v.push_back(std::move(th));  
    }  
    for(auto& t : v) t.join();  
}
```

Ошибка!!! `mastr_id` может не  
получить значение до вызова  
`bt()`

```
admin@irbis8:~/exp/lectures$ ./ex6  
Hello from master thread  
Hello  
Hello  
Hello  
Hello  
Hello  
Hello  
Hello  
Hello
```



# Вычисление интеграла

```
void integrate(int n, int start, int step, const
std::function <double (double)>& f, double&
rv) {
    double h = (B - A) / n;
    double v = 0;
    for(int i = start; i < n; i += step) {
        double x = A + i*h + h/2;
        v += f(x);
    }
    rv = v * h;
}
```



# Вычисление интеграла

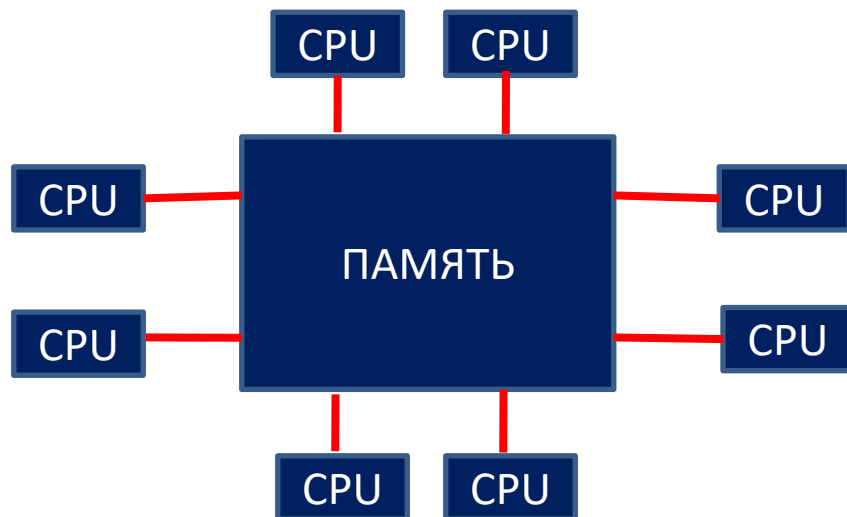
```
int main(int argc, char* argv[]) {  
    int n, nt;  
    if(argc != 3) {  
        std::cout << "usage: " << argv[0] << "  
number_of_boxes number_of_threads\n";  
        return -1;  
    }  
    n = atoi(argv[1]);  
    nt = atoi(argv[2]);  
    double v[nt];  
    std::thread thr[nt];
```

```
    for(int i = 0; i < nt; i++) {  
        std::thread t(integrate, n, i, nt,  
std::cref(sin),std::ref(v[i]));  
        thr[i] = std::move(t);  
    }  
    double u = 0;  
    for(int i = 0; i < nt; i++) {  
        thr[i].join();  
        u += v[i];  
    }  
    std::cout << "The integral value is " << u << "  
\n";  
    return 0;  
}
```



# Экспериментальные исследования

Эксперименты проводились на с двумя процессорами Intel®Xeon® E5620 (2.40)ГГц, 16 Гб оперативной памяти, 8 вычислительных ядер суммарно с гипертрейдингом – 16 виртуальных.



# Результат работы программы на экране

```
mposypkin@u5: ~/exp/myanm/numint
mposypkin@u5:~/exp/myanm/numint$ time ./numintp 100000000 1
The integral value is 2

real    0m11.088s
user    0m11.064s
sys     0m0.024s
mposypkin@u5:~/exp/myanm/numint$ time ./numintp 100000000 2
The integral value is 2

real    0m5.557s
user    0m11.067s
sys     0m0.030s
mposypkin@u5:~/exp/myanm/numint$ time ./numintp 100000000 4
The integral value is 2

real    0m2.788s
user    0m11.072s
sys     0m0.032s
mposypkin@u5:~/exp/myanm/numint$ time ./numintp 100000000 8
The integral value is 2

real    0m1.529s
user    0m11.330s
sys     0m0.040s
mposypkin@u5:~/exp/myanm/numint$ time ./numintp 100000000 16
The integral value is 2

real    0m1.073s
user    0m16.429s
sys     0m0.091s
mposypkin@u5:~/exp/myanm/numint$ time ./numintp 100000000 32
The integral value is 2

real    0m1.074s
user    0m16.465s
sys     0m0.116s
mposypkin@u5:~/exp/myanm/numint$
```

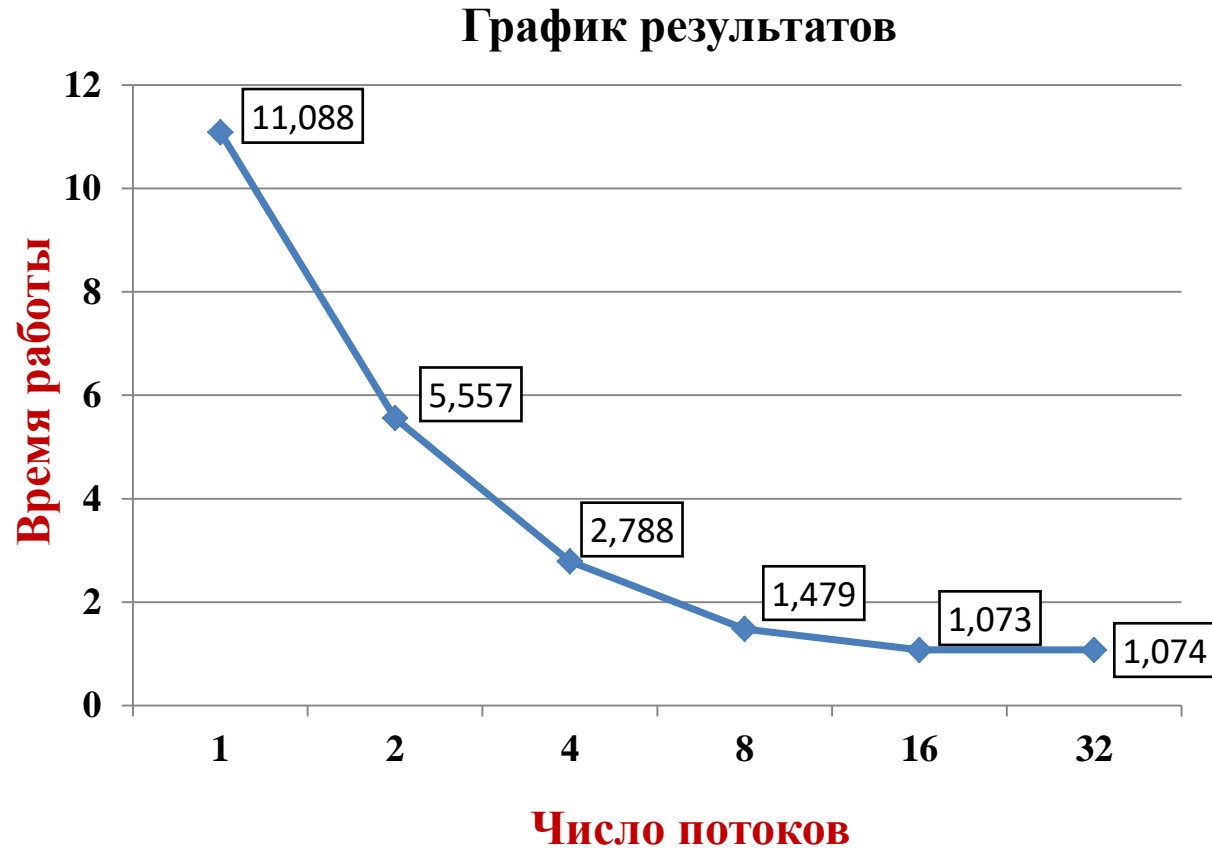


## Таблица результатов эксперимента время решения на языке C++11

Число потоков	1	2	4	8	16	32
Время работы	11.088	5.557	2.788	1.479	1.073	1.074



# График зависимости времени работы параллельной программы от числа запускаемых потоков на языке C++11



# Мьютексы

Класс **std::mutex**

Методы:

**lock()** – захват мьютекса (должен вызываться только если поток не владеет мьютексом) – приводит к блокировке потока до момента освобождения мьютекса другим потоком

**unlock()** – освобождение мьютекса (должен вызываться только, если мьютекс захвачен потоком)

**bool try\_lock()** – Пытается заблокировать мьютекс. Возврат происходит немедленно. В случае успешной установки блокировки возвращается true, в противном случае возвращается false.





# Resource Acquisition Is Initialization (RAII)

Концепция, предполагающая что получение ресурсов происходит при создании объекта, а освобождение происходит при вызове деструктора.

Это позволяет гарантировать освобождение в случае возникновения исключений, т.к. при обработке исключений вызываются деструкторы объектов, находящихся на стеке.



# Опасный подход к выделению ресурсов

```
class A {  
public:  
void alloc() { printf("resource allocated\n");}  
void free() { printf("resource freed\n"); }  
};
```

```
main() {  
    A a;
```

```
    try {  
        a.alloc();  
        throw E();  
        a.free();  
    } catch(...) {  
    }  
}
```

Не  
вызывается

```
mikdebian@debian32:~/exp/course/c11$ ./rail1  
resource allocated  
mikdebian@debian32:~/exp/course/c11$
```



# Правильный подход к управлению ресурсами

```
class A {  
public:  
    A() { printf("resource allocated\n");}  
    ~A() { printf("resource freed\n"); }  
};
```

```
struct E {  
};
```

```
main() {  
    try {  
        A a;  
        throw E();  
    } catch(...) {  
    }  
}
```

Деструктор вызывается  
при возникновении  
исключения и ресурсы  
освобождаются

```
mikdebian@debian32: ~/exp/course/c11$ ./raii  
resource allocated  
resource freed  
mikdebian@debian32: ~/exp/course/c11$
```



# RAII применительно к мьютексам

```
#include <mutex>
```

```
std::mutex mut;
```

```
int a;
```

```
void f() {
```

```
    std::lock_guard<std::mutex> lg(mut);
```

```
    a ++;
```

```
}
```

`std::lock_guard` – обертка для мьютекса, обеспечивающая его захват при создании «гарда» и освобождение на выходе из блока (в том числе и при исключении)



# Взаимная блокировка и ее преодоление

Thread1:

```
mutex1.lock();  
mutex2.lock();
```

Thread2:

```
mutex2.lock();  
mutex1.lock();
```

Необходимость захвата двух мьютексов может быть вызвана защитой двух элементов данных в некоторой операции (mutex1 защищает первый элемент, mutex2 – второй).



# Взаимная блокировка и ее преодоление

Thread1:

```
mutex1.lock();  
mutex2.lock();
```

Thread2:

```
mutex2.lock();  
mutex1.lock();
```

Возникает взаимная блокировка из-за того, что каждый из потоков пытается выполнить захват мьютекса, захваченного ранее другим потоком.



# Взаимная блокировка и ее преодоление

```
template< class Lockable1, class Lockable2, class LockableN... >  
void lock( Lockable1& lock1, Lockable2& lock2, LockableN& lockn... );
```

Выполняет безопасную блокировку нескольких мьютексов одновременно

Пример использования для защиты некоторой операции some\_op()

```
{  
std::lock(lhs.m,rhs.m);  
std::lock_guard<std::mutex> lock_a(lhs.m,std::adopt_lock);  
std::lock_guard<std::mutex> lock_b(rhs.m,std::adopt_lock);  
return some_op(lhs.some_detail, rhs.some_detail);  
}
```

std::adopt\_lock означает, что lock\_guard создается на базе заблокированного ранее мьютекса



# std::unique\_lock

Класс `unique_lock` является оболочкой над мьютексами, которая обладая всем потенциалом `lock_guard` позволяет выполнять отложенную блокировку, попытки блокировки с ограничением по времени ожидания, рекурсивную блокировку и использование с **условными переменными**.

The class `unique_lock` is movable, but not copyable -- it meets the requirements of `MoveConstructible` and `MoveAssignable` but not of `CopyConstructible` or `CopyAssignable`.



# Решение проблемы взаимной блокировки с использованием `unique_lock`

```
void transfer(Box &from, Box &to, int num)
{
    // don't actually take the locks yet
    std::unique_lock<std::mutex> lock1(from.m, std::defer_lock);
    std::unique_lock<std::mutex> lock2(to.m, std::defer_lock);

    // lock both unique_locks without deadlock
    std::lock(lock1, lock2);

    from.num_things -= num;
    to.num_things += num;

    // 'from.m' and 'to.m' mutexes unlocked in 'unique_lock' destructors
}
```

# Основные методы `unique_lock`

**`lock`** – блокирует ассоциированный мьютекс

**`try_lock`** выполняет попытку захвата мьютекса, возвращает управление, если мьютекс недоступен (захвачен другим потоком)

**`try_lock_for`** для `TimeLockable` мьютексов выполняет попытку захвата, возвращает управление, если мьютекс недоступен в течение указанного периода

**`try_lock_until`** для `TimeLockable` мьютексов выполняет попытку захвата, возвращает управление, если мьютекс недоступен до указанного момента времени

**`unlock`** разблокирует ассоциированный мьютекс

# shared\_mutex (C++ 17)

Подразумевает два типа владения: эксклюзивный и не эксклюзивный. В неэксклюзивном режиме мьютексом могут владеть сразу несколько потоков. Это дает возможность различать блокировку на чтение и запись.

Основные методы:

Exclusive locking

<b>lock</b>	locks the mutex, blocks if the mutex is not available
<b>try_lock</b>	tries to lock the mutex, returns if the mutex is not available
<b>unlock</b>	unlocks the mutex

Shared locking

<b>lock_shared</b>	locks the mutex for shared ownership, blocks if the mutex is not available
<b>try_lock_shared</b>	tries to lock the mutex for shared ownership, returns if the mutex is not available
<b>unlock_shared</b>	unlocks the mutex (shared ownership)

# shared\_lock (C++ 14)

Оболочка общего назначения для мьютексов, допускающих два режима блокировки (эксклюзивный режим и неэксклюзивный).

```
<template< class Mutex >  
class shared_lock;
```

Mutex - the type of the shared mutex to lock. The type must meet the SharedMutex requirements

SharedMutex:

std::shared\_mutex(since C++17)

std::shared\_timed\_mutex(since C++14)

```
class ThreadSafeCounter {  
public:  
    ThreadSafeCounter() = default;
```

// Multiple threads/readers can read the counter's value at the same time.

```
    unsigned int get() const {  
        std::shared_lock<std::shared_mutex> lock(mutex_);  
        return value_;  
    }
```

// Only one thread/writer can increment/write the counter's value.

```
    void increment() {  
        std::unique_lock<std::shared_mutex> lock(mutex_);  
        value_++;  
    }
```

// Only one thread/writer can reset/write the counter's value.

```
void reset() {  
    std::unique_lock<std::shared_mutex> lock(mutex_);  
    value_ = 0;  
}
```

private:

```
mutable std::shared_mutex mutex_;  
unsigned int value_ = 0;  
};
```

```
int main() {
    ThreadSafeCounter counter;

    auto increment_and_print = [&counter]() {
        for (int i = 0; i < 3; i++) {
            counter.increment();
            std::cout << std::this_thread::get_id() << ' ' << counter.get() << '\n';

            // Note: Writing to std::cout actually needs to be synchronized as well
            // by another std::mutex. This has been omitted to keep the example
            small.
        }
    };

    std::thread thread1(increment_and_print);
    std::thread thread2(increment_and_print);

    thread1.join();
    thread2.join();
}
```

# Условные переменные

Тип **std::condition\_variable**

Основные методы

**notify\_one** – уведомить один ждущий поток

**notify\_all** – уведомить все ждущие потоки

**wait** - блокировка потока до пробуждения

**wait\_for** – блокировка до пробуждения или истечения времени

**wait\_until** – блокировка до пробуждения или до заданного момента времени





# Ожидание «пробуждения»

```
void wait( std::unique_lock<std::mutex>& lock );
```

Ставит поток в очередь на ожидание по данной переменной. Возвращает управление в случае если другой поток вызвал **notify\_one ()** или **notify\_all()** по этой переменной. Допускаются также спонтанные срабатывания.

Для защиты от ложных срабатываний добавляется предикат

```
template< class Predicate >  
void wait( std::unique_lock<std::mutex>& lock, Predicate pred );
```

Сигнатура предиката: **bool pred()**

Эквивалентно коду:

```
while (!pred()) {  
    wait(lock);  
}
```



# Условные переменные: пример

```
std::mutex mut;
std::queue<data_chunk> data_queue;
std::condition_variable data_cond;
void data_preparation_thread()
{
    while(more_data_to_prepare())
    {
        data_chunk const data=prepare_data();
        std::lock_guard<std::mutex> lk(mut);
        data_queue.push(data);
        data_cond.notify_one();
    }
}
```



# Условные переменные: пример

```
void data_processing_thread()
{
    while(true)
    {
        std::unique_lock<std::mutex> lk(mut);
        data_cond.wait(lk, []{return !data_queue.empty();});
        data_chunk data=data_queue.front();
        data_queue.pop();
        lk.unlock();
        process(data);
        if(is_last_chunk(data))
            break;
    }
}
```

Лямбда-  
выражение



# Модификатор **mutable**

Переменная или поле данных класса, объявленное с этим модификатором, может изменяться в константных функциях.

Позволяет делать константные методы thread-safe, защищая чтение данных с помощью mutex-ов, объявленных с помощью mutable.



# Многопоточная очередь

```
template<typename T>
class threadsafe_queue
{
private:
mutable std::mutex mut;
std::queue<T> data_queue;
std::condition_variable data_cond;
```

```
public:
threadsafe_queue() {}
```

```
threadsafe_queue(threadsafe_queue const&
other)
{
std::lock_guard<std::mutex> lk(other.mut);
data_queue=other.data_queue;
}
```

mutable нужен для  
использования мьютекса в  
константных методах



# Многопоточная очередь

```
void push(T new_value)
{
    std::lock_guard<std::mutex> lk(mut);
    data_queue.push(new_value);
    data_cond.notify_one();
}
```

```
void wait_and_pop(T& value)
{
    std::unique_lock<std::mutex> lk(mut);
    data_cond.wait(lk,[this]{return !data_queue.empty();});
    value=data_queue.front();
    data_queue.pop();
}
```



# Многопоточная очередь

```
std::shared_ptr<T> wait_and_pop()
{
    std::unique_lock<std::mutex> lk(mut);
    data_cond.wait(lk,[this]{return !data_queue.empty();});
    std::shared_ptr<T>
    res(std::make_shared<T>(data_queue.front()));
    data_queue.pop();
    return res;
}
```

В константном методе  
используется mutex,  
объявленный как  
mutable

```
bool empty() const
{
    std::lock_guard<std::mutex> lk(mut);
    return data_queue.empty();
}
```



# Атомарные типы

- Над данными атомарного типа определены атомарные операции, т.е. операции, состояние которых доступно другим потокам либо до, либо после, но не в процессе операции.
- Атомарные операции могут читаться/записываться без синхронизации разными потоками.
- Одновременное обращения к общим неатомарным данным без синхронизации приводит к неопределенному поведению.



# Общий шаблон для атомарных ТИПОВ

**template< class T > struct atomic;**

Определяет атомарную «версию» типов, которые:

1. поддерживают тривиальную генерируемую операцию присваивания;
2. поддерживают тривиальное побитовое сравнение.

# Пример

```
#include <iostream>
#include <atomic>
```

```
class A {
public:
    A(int a) : mA(a) {
    }

    virtual void f() = 0;

protected:
    int mA;
};
```

```
class B : public A {
public:
    B(int b) : A(b) {
    }

    void f() {
        std::cout << A::mA;
    }
};
```

```
class C {
public:
    C(int a) : mA(a) {
    }

    void f() {
        std::cout << mA << "\n";
    }

protected:
    int mA;
};
```



# Какие варианты правильные?

```
int main() {  
  
    B b(1);  
    A& a = b;  
    C c(3);  
    std::atomic<A> atomicb(a);  
    std::atomic<B> atomicb(b);  
    std::atomic<C> atomicc(c);  
    return 0;  
}
```

# Какие варианты правильные?

```
int main() {
```

```
    B b(1);
```

```
    A& a = b;
```

```
    C c(3);
```

```
    std::atomic<A> atomicb(a);
```

```
    std::atomic<B> atomicb(b);
```

```
    std::atomic<C> atomicc(c);
```

```
    return 0;
```

```
}
```

Абстрактный класс

Имеет виртуальную  
функцию f()

# Атомарные варианты для арифметических типов

<code>std::atomic_bool</code>	<code>std::atomic&lt;bool&gt;</code>
<code>std::atomic_char</code>	<code>std::atomic&lt;char&gt;</code>
<code>std::atomic_schar</code>	<code>std::atomic&lt;signed char&gt;</code>
<code>std::atomic_uchar</code>	<code>std::atomic&lt;unsigned char&gt;</code>
<code>std::atomic_short</code>	<code>std::atomic&lt;short&gt;</code>
<code>std::atomic_ushort</code>	<code>std::atomic&lt;unsigned short&gt;</code>
<code>std::atomic_int</code>	<code>std::atomic&lt;int&gt;</code>
<code>std::atomic_uint</code>	<code>std::atomic&lt;unsigned int&gt;</code>
<code>std::atomic_long</code>	<code>std::atomic&lt;long&gt;</code>
<code>std::atomic_ulong</code>	<code>std::atomic&lt;unsigned long&gt;</code>
<code>std::atomic_llong</code>	<code>std::atomic&lt;long long&gt;</code>
<code>std::atomic_ullong</code>	<code>std::atomic&lt;unsigned long long&gt;</code>

`std::atomic<T*>` определен для всех типов указателей.

# Копирование атомарных объектов

Атомарные объекты нельзя копировать, поскольку копирование вовлекает чтение и запись двух объектов – такая операция не может быть атомарной.

# Копирование атомарных объектов

- Атомарные объекты нельзя присваивать друг другу и нельзя использовать в конструкторе копирования
- Объекты базового типа могут использоваться в конструкторе атомарной версии (конструктор не является атомарной операцией)
- Можно присваивать значения атомарных объектов соответствующим типам языка C++, эти операции атомарны
- Можно присваивать атомарному объекту значения базового типа, эта операция атомарна

# Примеры присваивания atomic

```
#include <iostream>
#include <atomic>
int main() {
    int a1 = 1;
    int a2 = 2;
    std::atomic<int> ai1;
    std::atomic<int> ai2(a2);
    ai1 = ai2;
    return 0;
}
```

```
./atomiccopy.cpp: In function 'int main()':
./atomiccopy.cpp:11:7: error: use of deleted function 'std::atomic<int>
& std::atomic<int>::operator=(const std::atomic<int>&)'
    ai1 = ai2;
    ^
In file included from ./atomiccopy.cpp:2:0:
/usr/include/c++/5/atomic:613:15: note: declared here
    atomic& operator=(const atomic&) = delete;
    ^
posypkin@mikhail:~/exp/sandbox/c11/atomic$
```



# Примеры присваивания atomic

```
#include <iostream>
```

```
#include <atomic>
```

```
int main() {  
    int a1 = 1;  
    int a2 = 2;  
    std::atomic<int> ai1;  
    std::atomic<int> ai2(a2);  
    ai1 = a1;  
    a2 = ai1;  
    std::cout << "a1 = " << a1 << ", a2 = " << a2 << ", ai1 = " << ai1 << ", ai2 = " << ai2 <<  
    "\n";  
    return 0;  
}
```

```
posypkin@mikhail:~/exp/sandbox/c11/atomic$ ./atomiccopy  
a1 = 1, a2 = 1, ai1 = 1, ai2 = 2
```

# Различия со стандартным присваиванием

- Обычное присваивание возвращает ссылку на объект из левой части

`T& class_name :: operator= ( T )`

- Присваивание атомарной переменной возвращает присаиваемое значение

`T operator=( T desired );`

# Пример, показывающий различия в семантике присваивания

```
#include <iostream>
```

```
#include <atomic>
```

```
int main() {
```

```
    int a = 1;
```

```
    (a = 2) = 3;
```

Ok

```
    std::cout << "a = " << a << "\n";
```

Напечатает 3

```
    std::atomic<int> ai(a);
```

```
    ai = 5;
```

Ok

```
    (ai = a) = 5;
```

```
    return 0;
```

```
}
```

Ошибка (нельзя присваивать не l-value)

# Lock Free

- Для некоторых типов реализация `std::atomic` не требует использования внешних синхронизаций, так как на целевой платформе поддерживается аппаратная синхронизация для базовых типов (`int`, `double` ...). (**lock free**)
- Для других типов, напротив, требуется использование примитивов синхронизации (мьютексы и т.п.). (**not lock free**)
- Аппаратная поддержка атомарности операций намного эффективнее, так как не требует синхронизации.

# Проверка lock-free

- Метод **is\_lock\_free** возвращает **true** для безблокировочных атомарных типов и **false** для всех остальных
- Стандартные атомарные типы, как правило, **lock free**

# Пример

```
#include <iostream>
#include <atomic>
```

```
struct C1 {
    int mA;
    double mD;
};
```

```
struct C2 {
    char mC[100];
};
```

```
struct C3 {
    int mA;
    double mD;
    double mH;
};
```

```
struct C4 {
    int mA;
    double mD;
    double mH;
    char mC[100];
};
```

# Пример

```
int main() {  
  
    std::atomic_int ai;  
    std::cout << "atomic_int is " << (ai.is_lock_free() ? "lock free " : "not  
lock free ") << "\n";  
  
    std::atomic<int*> api;  
    std::cout << "std::atomic<int*> is " << (api.is_lock_free() ? "lock free  
" : "not lock free ") << "\n";  
}
```

# Пример

```
{  
    C1 c;  
    std::atomic<C1> atomicc(c);  
    std::cout << "atomic<C1> is " << (atomicc.is_lock_free() ? "lock free  
" : "not lock free ") << "\n";  
}
```

```
{  
    C2 c;  
    std::atomic<C2> atomicc(c);  
    std::cout << "atomic<C2> is " << (atomicc.is_lock_free() ? "lock free  
" : "not lock free ") << "\n";  
}
```



# Пример

```
{
    C3 c;
    std::atomic<C3> atomicc(c);
    std::cout << "atomic<C3> is " << (atomicc.is_lock_free() ? "lock free " :
    "not lock free ") << "\n";
}

{
    C4 c;
    std::atomic<C4> atomicc(c);
    std::cout << "atomic<C4> is " << (atomicc.is_lock_free() ? "lock free " :
    "not lock free ") << "\n";
}

return 0;
}
```

# Результат выполнения

В gcc необходимо при сборке подключать библиотеку `atomic` с помощью флага **-latomic**

```
posypkin@mikhail: ~/exp/sandbox/c11/atomic
posypkin@mikhail:~/exp/sandbox/c11/atomic$ c++ -std=c++11 -o lockfree ./lockfree.cpp -latomic
posypkin@mikhail:~/exp/sandbox/c11/atomic$ ./lockfree
atomic_int is lock free
std::atomic<int*> is lock free
atomic<C1> is lock free
atomic<C2> is not lock free
atomic<C3> is not lock free
atomic<C4> is not lock free
posypkin@mikhail:~/exp/sandbox/c11/atomic$
```

# std::atomic\_flag

- Простейший атомарный объект логического (булевского) типа.
- Является lock free
- Не допускает присваивания и конструктора копирования

# Методы `std::atomic`

- **clear** – атомарная операция присваивания флагу значения `false`
- **test\_and\_set** – атомарная операция, которая устанавливается значение флага в `true` и возвращает его предыдущее значение

# vector::emplace\_back

- Создание элемента контейнера «на месте» без дополнительного копирования
- Как правило, выигрывает по производительности у push\_back

`template< class... Args > void emplace_back( Args&&... args );` C++ 11


`template< class... Args > reference emplace_back( Args&&... args );` C++ 17  
(возвращает ссылку на созданный элемент в контейнере)

# Пример использования atomic\_flag для организации блокировок

```
#include <unistd.h>
#include <thread>
#include <vector>
#include <iostream>
#include <atomic>

std::atomic_flag lock = ATOMIC_FLAG_INIT;

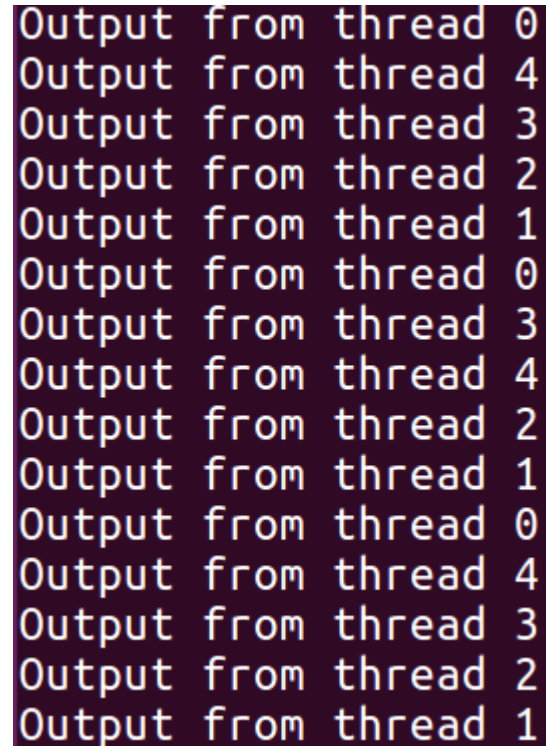
void f(int n)
{
    for (int cnt = 0; cnt < 10; ++cnt) {
        while (lock.test_and_set()) ;
        std::cout << "Output from thread " << n << '\n';
        lock.clear();
        sleep(1);
    }
}
```



Циклическая  
блокировка (не  
лучший, хотя и  
корректный,  
способ)

# Пример использования `atomic_flag` для организации блокировок

```
int main()
{
    std::vector<std::thread> v;
    for (int n = 0; n < 5; ++n) {
        v.emplace_back(f, n);
    }
    for (auto& t : v) {
        t.join();
    }
}
```



```
Output from thread 0
Output from thread 4
Output from thread 3
Output from thread 2
Output from thread 1
Output from thread 0
Output from thread 3
Output from thread 4
Output from thread 2
Output from thread 1
Output from thread 0
Output from thread 4
Output from thread 3
Output from thread 2
Output from thread 1
Output from thread 0
Output from thread 4
Output from thread 3
Output from thread 2
Output from thread 1
```

# Методы класса `std::atomic`

- `T load( std::memory_order order = std::memory_order_seq_cst ) const;` - атомарное чтение значения атомарного объекта (эквивалентна `()` )
- `void store( T desired, std::memory_order order = std::memory_order_seq_cst );` - атомарная запись данных в атомарный объект (эквивалентна `=` )



# Методы класса `std::atomic`

- `T exchange( T desired, std::memory_order order = std::memory_order_seq_cst );`

Атомарная операция. Заменяет значения атомарной переменной на значение `desired`, возвращает старое значение.

# Методы класса `std::atomic`

- `bool compare_exchange_weak( T& expected, T desired, std::memory_order order = std::memory_order_seq_cst );`
- `bool compare_exchange_strong( T& expected, T desired, std::memory_order order = std::memory_order_seq_cst );`

Если значение атомарной переменной совпадает со значением `expected`, то ее значение заменяется на `desired`. Возвращаемое значение при этом `true`. В противном случае, значение атомарной переменной записывается в `expected`. Возвращаемое значение при этом `false`.

# Методы класса `std::atomic`

- `bool compare_exchange_weak` – не всегда срабатывает, если `expected` совпадает со значением атомарной переменной, поэтому должна использоваться в цикле с проверкой возвращаемого значения (*spurious failure*)
- `bool compare_exchange_strong` – если `expected` совпадает со значением атомарной переменной, то запись производится

Если значение атомарной переменной совпадает со значением `expected`, то ее значение заменяется на `desired`. Возвращаемое значение при этом `true`. В противном случае, значение атомарной переменной записывается в `expected`. Возвращаемое значение при этом `false`.

# Цикл для compare\_exchange\_weak

```
expected = current.load();  
do desired = function(expected);  
while (!current.compare_exchange_weak(expected, desired));
```

# Семантика

## compare\_exchange\_strong/weak

```
if (memcmp(object, expected, sizeof(*object)) == 0)
    memcpy(object, &desired, sizeof(*object));
else
    memcpy(expected, object, sizeof(*object));
```

Может отличаться от сравнения с помощью == и приводить к парадоксальным результатам. Особенно актуально для нецелочисленной арифметики.

# Атомарные операции сложного присваивания

```
T fetch_add( T arg,  
             memory_order = std::memory_order_seq_cst );
```

Атомарно добавляет `arg` к значению, хранящемуся в атомарном объекте, возвращает «старое» значение. Эквивалентно оператору `+=` (с точностью до аргумента, определяющего дисциплину работы с памятью)

```
T fetch_sub( T arg,    memory_order = std::memory_order_seq_cst );
```

Атомарно вычитает `arg` из значения, хранящегося в атомарном объекте, возвращает «старое» значение. Эквивалентно оператору `-=` (с точностью до аргумента, определяющего дисциплину работы с памятью)

# Атомарные операции сложного присваивания

```
T fetch_and( T arg, memory_order = std::memory_order_seq_cst );
```

Атомарно выполняет побитовое И между аргументом и значением атомного объекта, изменяя значение атомарного объекта, и возвращает «старое» значение. Эквивалентно оператору `&=` (с точностью до аргумента, определяющего дисциплину работы с памятью)

```
T fetch_or( T arg, memory_order = std::memory_order_seq_cst );
```

Атомарно выполняет побитовое ИЛИ между аргументом и значением атомного объекта, изменяя значение атомарного объекта, и возвращает «старое» значение. Эквивалентно оператору `|=` (с точностью до аргумента, определяющего дисциплину работы с памятью)

# Атомарные операции сложного присваивания

```
T fetch_xor ( T arg, memory_order =  
std::memory_order_seq_cst );
```

Атомарно выполняет побитовое ИСКЛЮЧАЮЩЕЕ ИЛИ между аргументом и значением атомного объекта, изменяя значение атомарного объекта, и возвращает «старое» значение. Эквивалентно оператору `|=` (с точностью до аргумента, определяющего дисциплину работы с памятью)



Operation	atomic_flag	atomic<bool>	atomic<T*>	atomic<integral-type>	atomic<other-type>
test_and_set	✓				
clear	✓				
is_lock_free		✓	✓	✓	✓
load		✓	✓	✓	✓
store		✓	✓	✓	✓
exchange		✓	✓	✓	✓
compare_exchange_weak, compare_exchange_strong		✓	✓	✓	✓
fetch_add, +=			✓	✓	
fetch_sub, -=			✓	✓	
fetch_or,  =				✓	
fetch_and, &=				✓	
fetch_xor, ^=				✓	
++, --			✓	✓	

# Модели памяти

- sequentially consistent ordering  
(memory\_order\_seq\_cst)
- acquire-release ordering  
(memory\_order\_consume,  
memory\_order\_acquire,  
memory\_order\_release, and  
memory\_order\_acq\_rel)
- relaxed ordering (memory\_order\_relaxed)

# Порядок работы с памятью

- На одном потоке порядок работы с памятью определяется выполнением операций чтения и записи и выражается отношением **«выполнилось ранее»**.
- Взаимный порядок выполнения операций на разных потоках определяется синхронизацией и определяется отношением **«синхронизовано с»**.

# Отношение «выполнилось ранее»

Означает, что в одном потоке одна операция выполнялась до другой. Например, в следующем коде:

1. `a = 5;`
2. `b = a;`

операция 1 выполняется до 2. Здесь тоже:

1. `a = 5;`
2. `b = 10.`

А вот порядок вычисления аргументов функции при вызове не определен, поэтому между выражениями в аргументах не возникает отношения «sequenced before»:

`f(a = 10, b = 5);`

# Отношение синхронизации

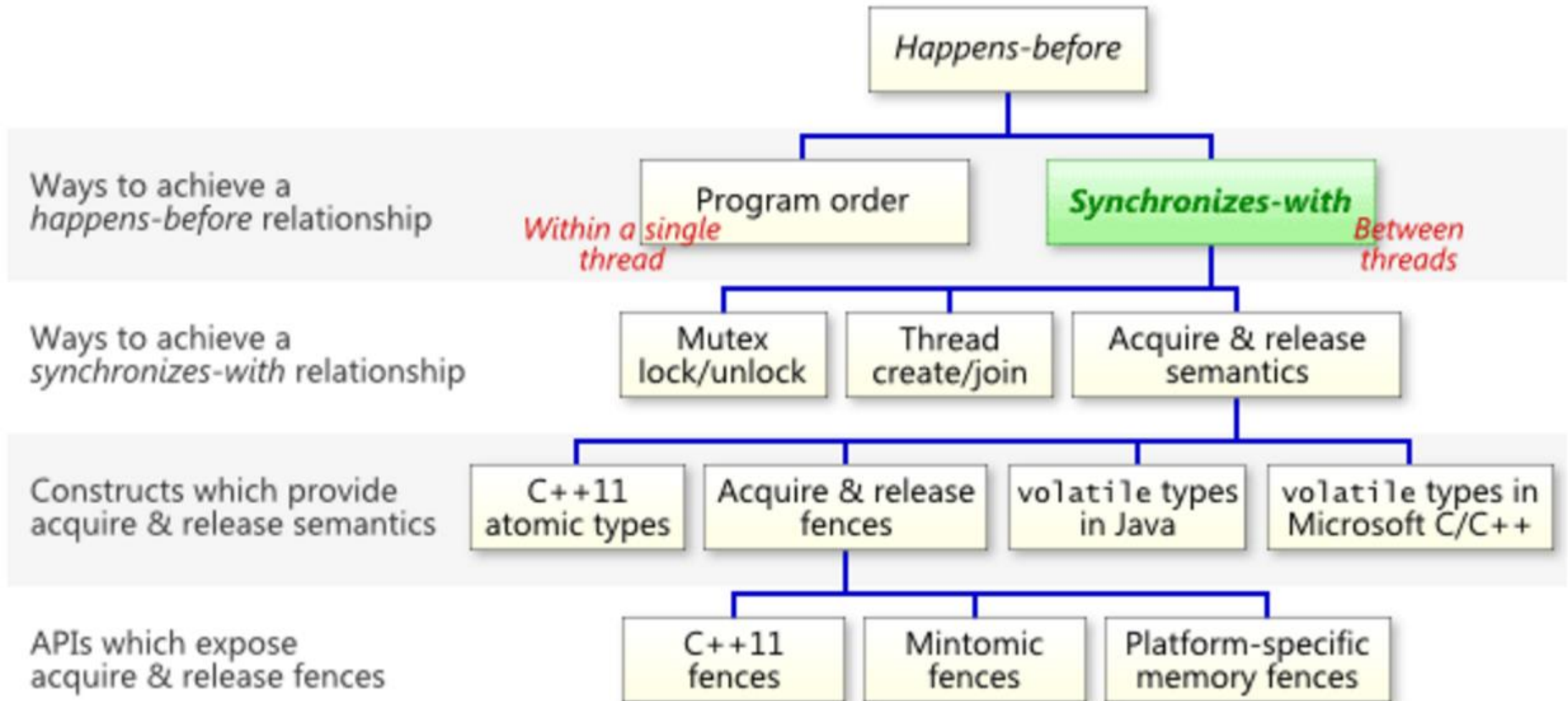
- Атомарная операция чтения  $R$  значения атомарной переменной (кроме **relaxed** и **consume**) синхронизируется с операцией  $W$  (кроме **relaxed**) записи этого значения на другом потоке, если прочитанное значение записано операцией  $W$  либо последующими операциями записи в том же потоке
- Операция блокировки мьютекса синхронизируется с операцией освобождения мьютекса.

# Отношение предшествования «случилось до»

- A happens before B, если выполнение A предшествует B в одном потоке («выполнилось ранее»)
- A happens before B, если A «синхронизовано с B» (A – запись, B – чтение)
- $A \text{ h.b. } B \text{ и } B \text{ h.b. } C \Rightarrow A \text{ h.b. } C$

**Выполнение программы должно соблюдать семантику отношения «случилось до». Операция видит все побочные эффекты, которые «случились до» нее.**

# Отношение «случилось до»



# Пример отношения предшествования

```
std::vector<int> data;  
std::atomic<bool> data_ready(false);
```

```
void reader_thread()  
{  
    while(!data_ready.load())  
    {  
        std::this_thread::sleep(std::milliseconds(1));  
    }  
    std::cout<<"The answer="<<data[0]<<"\n";  
}
```

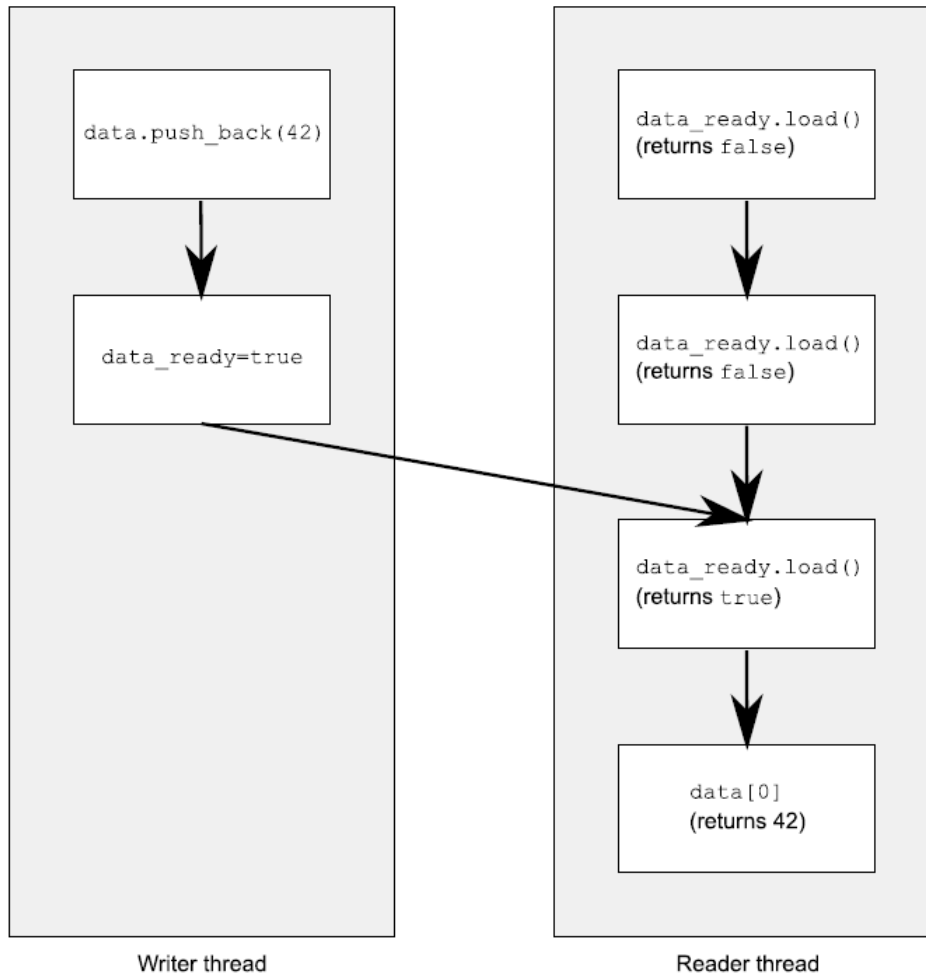
```
void writer_thread()  
{  
    data.push_back(42);  
    data_ready=true;  
}
```

← 42

Отношение предшествования, инициируемое операциями над атомарными переменными гарантирует, что прочитано будет значение 42, в вектор data произойдет после



# Граф отношения предшествования для данного примера



# Последовательная модель

- Все изменения данных одинаковым образом видны на всех потоках
- Изменения согласованы с отношением (случилось до)

# Пример последовательной модели

```
#include <atomic>
#include <thread>
#include <assert.h>
std::atomic<bool> x,y;
std::atomic<int> z;

void write_x()
{
    x.store(true, std::memory_order_seq_cst);
}

void write_y()
{
    y.store(true, std::memory_order_seq_cst);
}
```

# Пример последовательной модели

```
void read_x_then_y()  
{  
    while(!x.load(std::memory_order_seq_cst));  
    if(y.load(std::memory_order_seq_cst))  
        ++z;  
}
```

```
void read_y_then_x()  
{  
    while(!y.load(std::memory_order_seq_cst));  
    if(x.load(std::memory_order_seq_cst))  
        ++z;  
}
```

# Пример последовательной модели

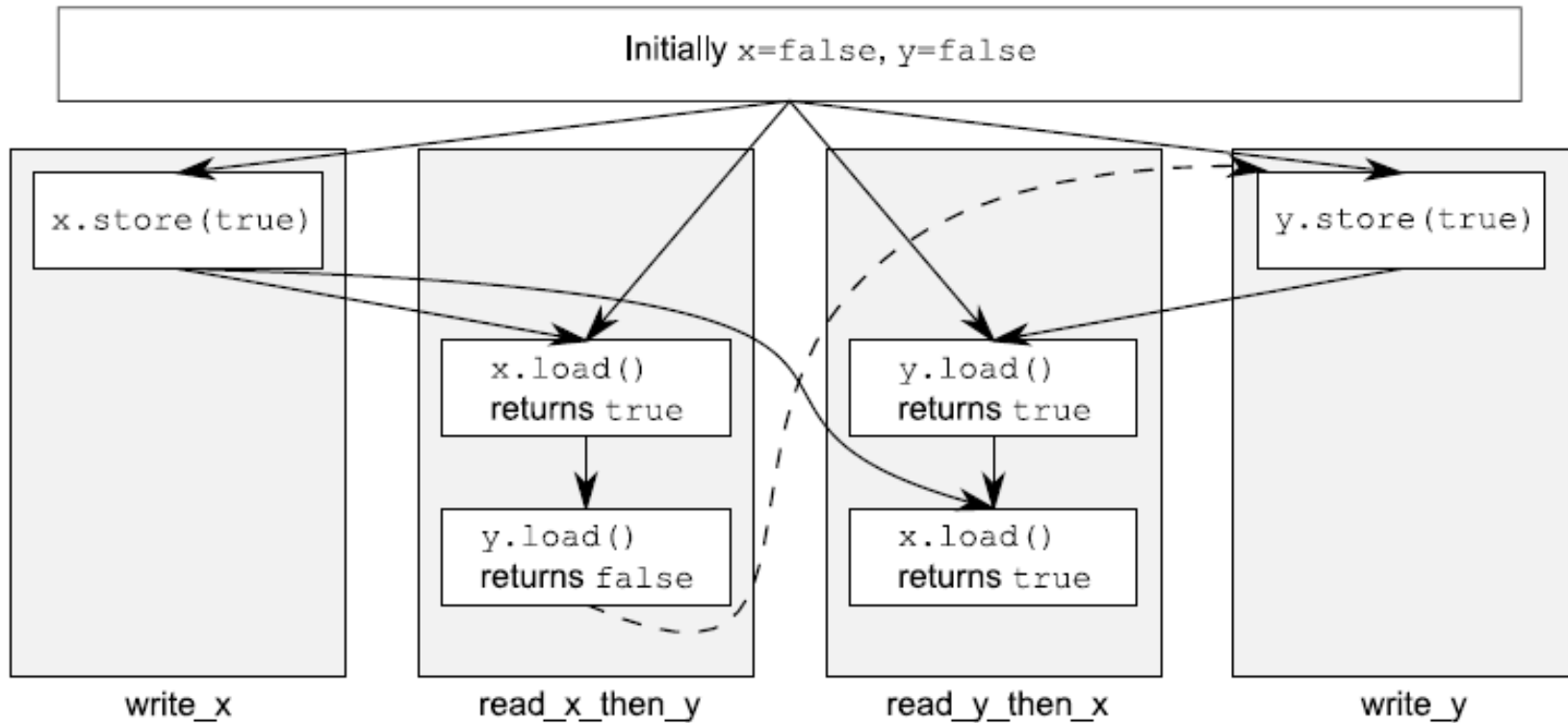
```
int main()
{
    x=false;
    y=false;
    z=0;
    std::thread a(write_x);
    std::thread b(write_y);
    std::thread c(read_x_then_y);
    std::thread d(read_y_then_x);
    a.join();
    b.join();
    c.join();
    d.join();
    assert(z.load()!=0);
}
```

# Объяснение

Обе переменных  $x$  или  $y$  получают значение `true`. Если в потоке, выполняющем `read_x_then_y`, операция **`y.load`** вернула **`false`**, то значит  $y$  примет значение **`true`** не раньше  $x$ . Поэтому в потоке, выполняющем **`read_y_then_x`**, операция **`x.load`** вернет **`true`**. Это обусловлено тем, что все потоки видят ситуацию одинаковым образом.

В последовательной модели предполагается, что есть некоторый порядок, в котором операции меняют состояние памяти и различные потоки наблюдают эти события одинаковым образом.

# Граф отношений предшествования



# Семантика Acquire-Release

- **Acquire** всегда применяется только для операций чтения, либо чтения-записи-модификации
- **Release** всегда применяется для операций записи, либо для чтения-записи-модификации



**Acquire semantics** A load operation with this memory order performs the acquire operation on the affected memory location: no reads or writes in the current thread can be reordered before this load. All writes in other threads that release the same atomic variable are visible in the current thread.

**Release semantics** A store operation with this memory order performs the release operation: no reads or writes in the current thread can be reordered after this store. All writes in the current thread are visible in other threads that acquire the same atomic variable (see Release-Acquire ordering below) and writes that carry a dependency into the atomic variable become visible in other threads that consume the same atomic.

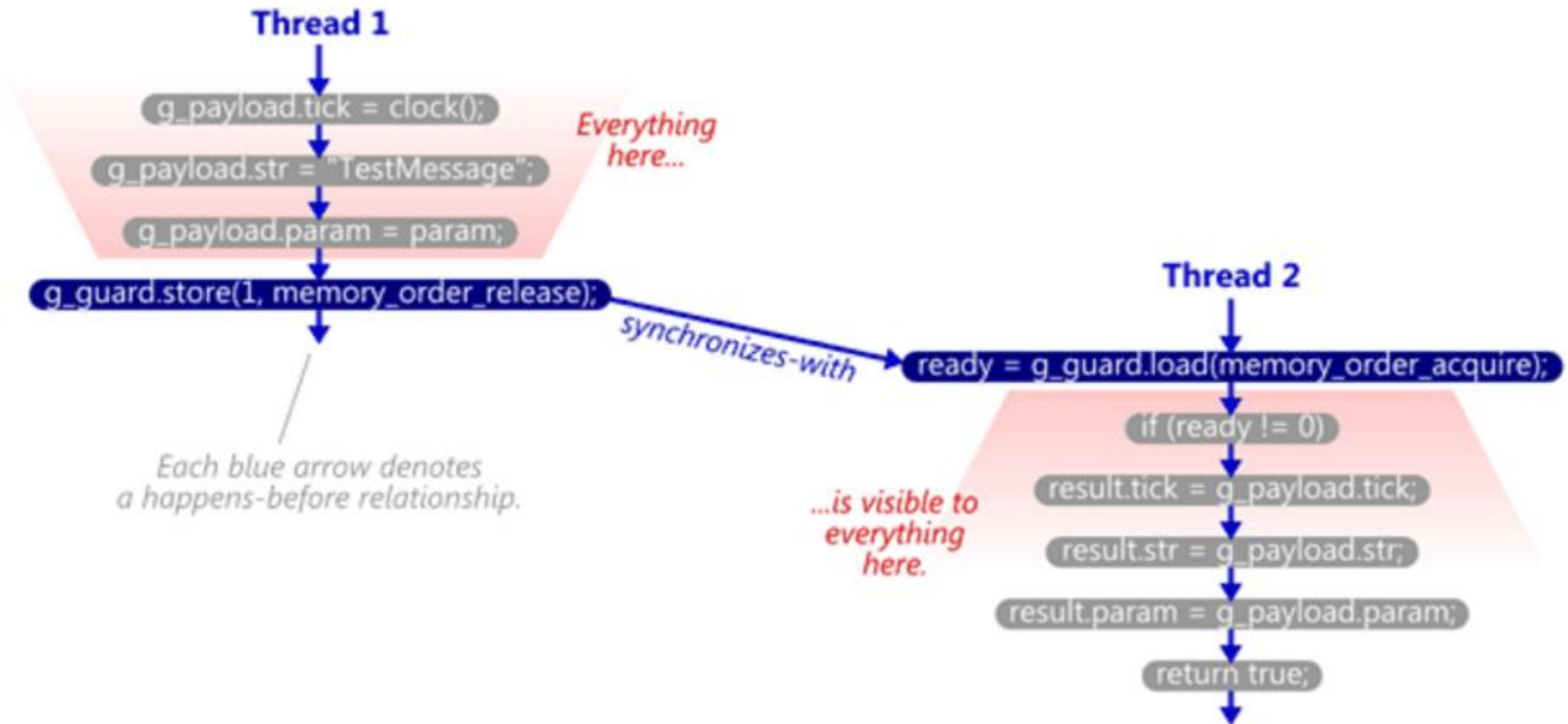
# Семантика Acquire-Release

- **Acquire** не дает компилятору/ран-тайму переместить операции чтения и записи до нее
- **Release** не дает компилятору/ран-тайму переместить операции чтения и записи после нее
- Поток выполняющий **acquire** «видит» изменения, произошедшие с общими переменными до **release**

# Использование acquire-release для «обмена» сообщениями

```
struct Message { clock_t tick; const char* str; void* param; };  
Message g_payload;  
std::atomic<int> g_guard(0);  
  
void SendTestMessage(void* param)  
{  
    // Copy to shared memory using non-atomic stores.  
    g_payload.tick = clock();  
    g_payload.str = "TestMessage";  
    g_payload.param = param;  
  
    // Perform an atomic write-release to indicate that the message is ready.  
    g_guard.store(1, std::memory_order_release);  
}
```

# Иллюстрация acquire-release



# Использование acquire-release для «обмена» сообщениями

```
std::atomic<int> g_guard(0);
```

```
void SendTestMessage(void* param)
```

```
{
```

```
    // Copy to shared memory using non-atomic stores.
```

```
    g_payload.tick = clock();
```

```
    g_payload.str = "TestMessage";
```

```
    g_payload.param = param;
```

```
    // Perform an atomic write-release to indicate that the message is ready.
```

```
    g_guard.store(1, std::memory_order_release);
```

```
}
```

# Использование acquire-release для «обмена» сообщениями

```
bool TryReceiveMessage(Message& result)
{
    // Perform an atomic read-acquire to check whether the message is ready.
    int ready = g_guard.load(std::memory_order_acquire);

    if (ready != 0)
    {
        // Yes. Copy from shared memory using non-atomic loads.
        result.tick = g_payload.tick;
        result.str  = g_payload.str;
        result.param = g_payload.param;

        return true;
    }
    return false;
}
```

# Пример использования Acquire-Release

```
#include <atomic>
#include <thread>
#include <assert.h>
std::atomic<bool> x,y;
std::atomic<int> z;
void write_x()
{
x.store(true,std::memory_order_release);
}
void write_y()
{
y.store(true,std::memory_order_release);
}
```

# Пример ограничений Acquire-Release

```
void read_x_then_y()  
{  
  while(!x.load(std::memory_order_acquire));  
  if(y.load(std::memory_order_acquire))  
    ++z;  
}
```

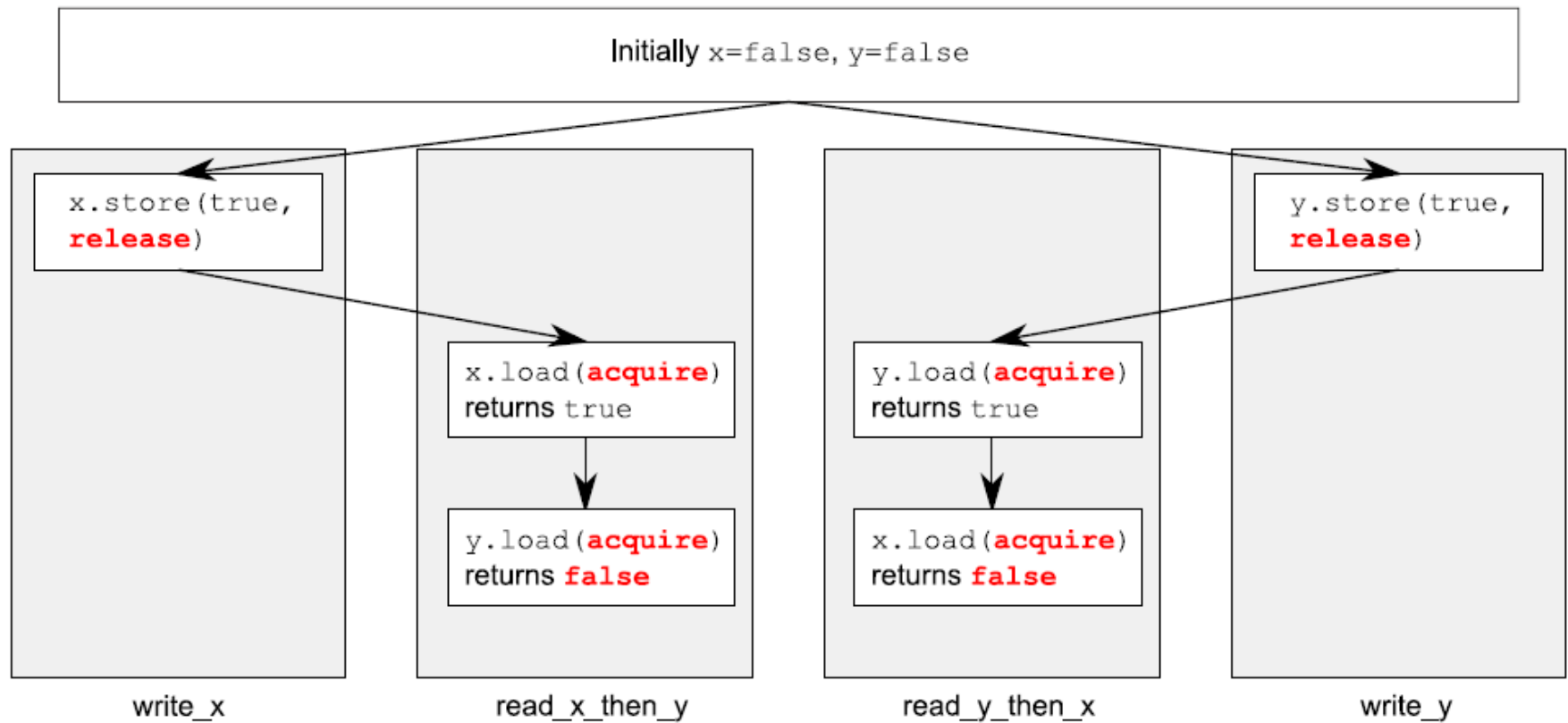
```
void read_y_then_x()  
{  
  while(!y.load(std::memory_order_acquire));  
  if(x.load(std::memory_order_acquire))  
    ++z;  
}
```



# Пример использования Acquire-Release

```
int main()
{
    x=false;
    y=false;
    z=0;
    std::thread a(write_x);
    std::thread b(write_y);
    std::thread c(read_x_then_y);
    std::thread d(read_y_then_x);
    a.join();
    b.join();
    c.join();
    d.join();
    assert(z.load()!=0);
}
```

**В данном  
случае  
срабатывание  
assert не  
гарантируется!**



**Figure 5.6** Acquire-release and happens-before

# Relaxed Model

```
#include <atomic>
#include <thread>
#include <assert.h>
std::atomic<bool> x,y;
std::atomic<int> z;

void write_x_then_y()
{
    x.store(true,std::memory_order_relaxed);
    y.store(true,std::memory_order_relaxed);
}
```

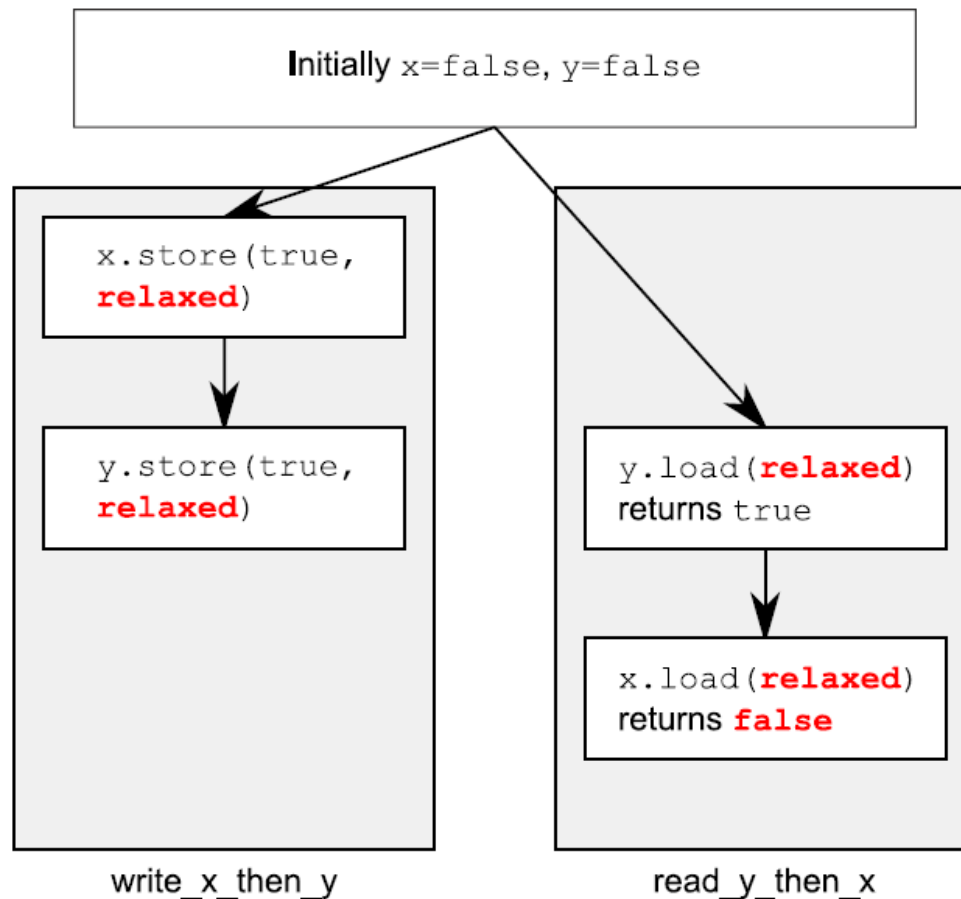
# Relaxed Model

```
void read_y_then_x()
{
    while(!y.load(std::memory_order_relaxed));
    if(x.load(std::memory_order_relaxed))
        ++z;
}

int main()
{
    x=false; y=false; z=0;
    std::thread a(write_x_then_y);
    std::thread b(read_y_then_x);
    a.join();
    b.join();
    assert(z.load()!=0);
}
```

# Relaxed Model

Проверка может сработать, т.к. релаксационная модель не гарантирует упорядочивания.



# Использование Relaxed Model

Relaxed model  
полезно  
использовать там,  
где не важен  
порядок изменения  
данных, но требуется  
атомарность.  
Например –  
счетчики.

```
1  // relaxed.cpp
2
3  #include <vector>
4  #include <iostream>
5  #include <thread>
6  #include <atomic>
7
8  std::atomic<int> cnt = {0};
9
10 void f()
11 {
12     for (int n = 0; n < 1000; ++n) {
13         cnt.fetch_add(1, std::memory_order_relaxed);
14     }
15 }
16
17 int main()
18 {
19     std::vector<std::thread> v;
20     for (int n = 0; n < 10; ++n) {
21         v.emplace_back(f);
22     }
23     for (auto& t : v) {
24         t.join();
25     }
26     std::cout << "Final counter value is " << cnt << '\n';
27 }
```

# Барьеры памяти

```
#include <atomic>
#include <thread>
#include <assert.h>
std::atomic<bool> x,y;
std::atomic<int> z;

void write_x_then_y()
{
    x.store(true,std::memory_order_relaxed);
    std::atomic_thread_fence(std::memory_order_release);
    y.store(true,std::memory_order_relaxed);
}
```

# Барьеры памяти

```
void read_y_then_x()
{
    while(!y.load(std::memory_order_relaxed));
    std::atomic_thread_fence(std::memory_order_acquire);
    if(x.load(std::memory_order_relaxed))
        ++ z;
}
```

```
int main()
{
    x=false;
    y=false;
    z=0;
    std::thread a(write_x_then_y);
    std::thread b(read_y_then_x);
    a.join();
    b.join();
    assert(z.load()!=0);
}
```



# Барьеры памяти

assert сработает, т.к. операция

```
y.store(true,std::memory_order_relaxed);
```

расположенная после release-барьера записала значение, прочитанное операцией

```
y.load(std::memory_order_relaxed),
```

расположенной до acquire-барьера, таким образом вызвав синхронизацию барьеров.

# Общее правило

## **Fence-fence synchronization**

A release fence FA in thread A synchronizes-with an acquire fence FB in thread B, if

- There exists an atomic object M,
- There exists an atomic write X (with any memory order) that modifies M in thread A
- FA is sequenced-before X in thread A
- There exists an atomic read Y (with any memory order) in thread B
- Y reads the value written by X (or the value would be written by release sequence headed by X if X were a release operation)
- Y is sequenced-before FB in thread B

In this case, all non-atomic and relaxed atomic stores that happen-before FA in thread A will be synchronized-with all non-atomic and relaxed atomic loads from the same locations made in thread B after FB.

# Синхронизация барьера и атомарного чтения

## **Fence-atomic synchronization**

A release fence F in thread A synchronizes-with atomic acquire operation Y in thread B, if

- there exists an atomic store X (with any memory order)
- Y reads the value written by X (or the value would be written by release sequence headed by X if X were a release operation)
- F is sequenced-before X in thread A

In this case, all non-atomic and relaxed atomic stores that happen-before X in thread A will be synchronized-with all non-atomic and relaxed atomic loads from the same locations made in thread B after F.

# Синхронизация атомарной записи и барьера

## **Atomic-fence synchronization**

An atomic release operation X in thread A synchronizes-with an acquire fence F in thread B, if

- there exists an atomic read Y (with any memory order)
- Y reads the value written by X (or by the release sequence headed by X)
- Y is sequenced-before F in thread B

In this case, all non-atomic and relaxed atomic stores that happen-before X in thread A will be synchronized-with all non-atomic and relaxed atomic loads from the same locations made in thread B after F.

# Использование барьеров для упорядочивания неатомарных операций с памятью

```
void SendTestMessage(void* param)
{
    g_payload.tick = clock();
    g_payload.str = "TestMessage";
    g_payload.param = param;

    std::atomic_thread_fence(std::memory_order_release);

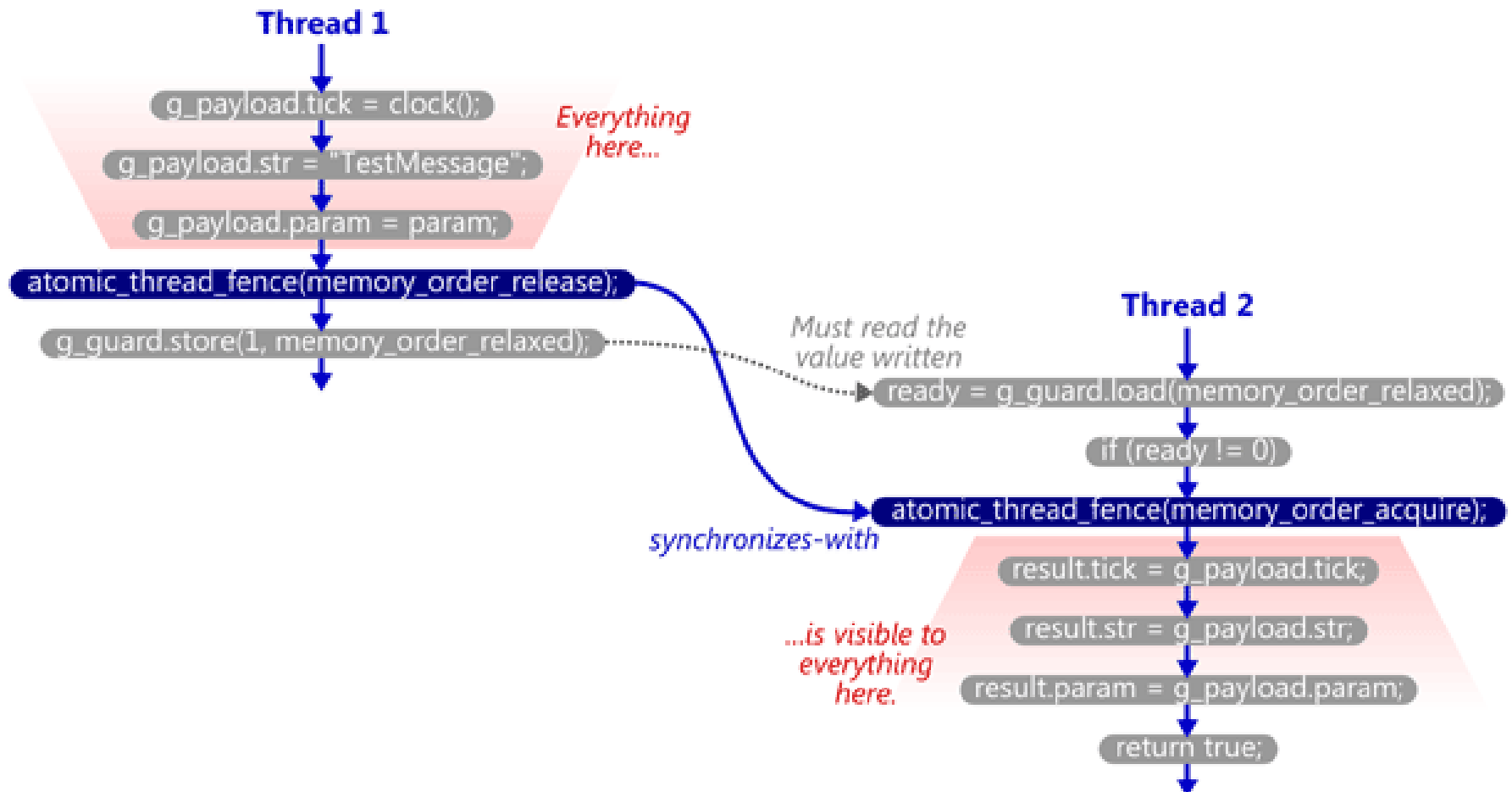
    g_guard.store(1, std::memory_order_relaxed);
}
```

# Использование барьеров для упорядочивания неатомарных операций с памятью

```
bool TryReceiveMessage(Message& result)
{
    int ready = g_guard.load(std::memory_order_relaxed);
    if (ready != 0)
    {
        std::atomic_thread_fence(std::memory_order_acquire);
        result.tick = g_payload.tick;
        result.str = g_msg_str;
        result.param = g_payload.param;

        return true;
    }
    return false;
}
```

# Использование барьеров для упорядочивания неатомарных операций с памятью



# std::future

- Основной инструмент для работы с результатом, возвращаемым функцией потока
- Позволяет идентифицировать событие, которое должно произойти в будущем и его результат
- Параметр шаблона T указывает на тип результата события `std::future<T>`
- `std::future<void>` означает что событие не связано с конкретным типом данных



# Конструкторы `std::future`

- **`future()`** - пустой конструктор, обычно не используется напрямую;
- **`future( future&& other )`** - конструктор перемещения;
- **`future( const future& other ) = delete`** - копирующий конструктор не поддерживается.

# Присваивание std::future

- **future& operator=( future&& other );** - перемещение объекта;
- **future& operator=( const future& other ) = delete;** - обычное копирование запрещено.

# Основные методы `std::future`

- `void wait() const;` - блокирует поток до момента, когда результат `std::future` готов
- `T get();` - возвращает значение результата;
- `T& get();` - ссылка на результат (для `std::future<T&>`;
- `void get()` – только для `std::future<void>`

# Основные методы `std::future`

- `template< class Rep, class Period > std::future_status  
wait_for( const std::chrono::duration<Rep,Period>&  
timeout_duration ) const;`

Ожидает получения результата не более заданного времени;

Возвращаемое значение:

**future\_status::deferred**      функция, которая должна вернуть результат, еще не стартовала;

**future\_status::ready**    результат готов;

**future\_status::timeout**      закончился период ожидания.

# Основные методы `std::future`

- `template< class Clock, class Duration > std::future_status  
wait_until( const std::chrono::time_point<Clock,Duration>&  
timeout_time ) const;`

Ожидает получения результата не позднее времени;

Возвращаемое значение:

**future\_status::deferred**          функция, которая должна вернуть результат, еще не стартовала;

**future\_status::ready**   результат готов;

**future\_status::timeout**        закончился период ожидания.

# Асинхронные операции

- **std::async** позволяет инициировать выполнение операции, результат которой может быть получен позднее
- Операцией является объект или функция (Callable object)
- Результатом является объект типа **std::future**

# Асинхронные операции

```
enum class launch  
{  
    async, deferred  
};
```

```
template<typename Callable,typename ... Args>  
future<result_of<Callable(Args...)>::type>  
async(Callable&& func,Args&& ... args);
```

```
template<typename Callable,typename ... Args>  
future<result_of<Callable(Args...)>::type>  
async(launch policy, Callable&& func,Args&& ... args);
```



# Пример использования

```
#include <future>
#include <iostream>

int find_the_answer_to_ltuae();

void do_other_stuff();

int main()
{
    std::future<int> the_answer=std::async(find_the_answer_to_ltuae);

    do_other_stuff();

    std::cout<<"The answer is "<<the_answer.get()<<std::endl;
}
```



# Пример использования

```
#include <thread>
#include <future>
#include <iostream>
#include <functional>

class A {
public:
    A(int a) : mA(a) { }

    int operator () (int b) const { return mA + b; }

    int mul(int b) const { return mA * b; }

    void update(int v) { mA = v; }

private:
    int mA;
};
```

# Пример использования

```
int main() {  
    A a(5);  
    std::future<int> fut = std::async(a, 4);  
    std::cout << "sum = " << fut.get() << "\n";  
    std::future<int> fut1 = std::async(&A::mul, a, 4);  
    std::cout << "mul = " << fut1.get() << "\n";  
  
    std::future<void> fut2 = std::async(&A::update, a, 10);  
    fut2.wait();  
    std::cout << "new sum = " << a(20) << "\n";  
  
    std::future<void> fut3 = std::async(&A::update, &a, 10);  
    fut3.wait();  
    std::cout << "new sum = " << a(20) << "\n";  
  
    std::future<int> fut4 = std::async([](int a, int b) { return a + b;}, 10, 100);  
    std::cout << "10 + 100 = " << fut4.get() << "\n";  
    return 0;  
}
```

# Пример использования

```
posypkin@mikhail:~/exp/sandbox/c11/async$ ./async1  
sum = 9  
mul = 20  
new sum = 25  
new sum = 30  
10 + 100 = 110
```

# Thread Oversubscription

**Oversubscription** - ситуация, когда число потоков превышает число физически доступных ядер.

Небольшое превышение может позитивно сказаться на производительности: когда один поток ожидает на операциях с памятью, другой выполняет вычисления. Это позволяет задействовать внутрипроцессорный параллелизм и, тем самым, ускорить выполнение программы.

# Негативные последствия Oversubscription: переключение контекста

Переключение контекста потока приводит к непродуктивным накладным расходам. Поэтому если число потоков намного превышает число ядер, то переключение контекста может существенно снизить производительность программы.

# Негативные последствия Oversubscription: cache cooling

Разные потоки, как правило, работают с разными данными. Поэтому при переключении ядра между потоками из кэш-памяти вытесняются данные, с которыми работал поток. Когда поток опять ставится на выполнение, увеличивается время на доступ к данным, т.к. требуется их загрузка в кэш из основной памяти.

# std::thread vs std::async

## std::thread

- + Больше контроля за выполнением потоков («ручное управление»)
- + Доступ к низкоуровневому API (pthreads)
- Проблемы балансировки и перегрузки системы (oversubscription)
- Сложный доступ к результату выполнения

## std::async

- + больше свободы системе в плане балансировки нагрузки (система решает запускать новый поток или нет)
- + легкий доступ к результату через std::future
- отсутствие низкоуровневого контроля

# Управление дисциплиной вызова

```
auto f6=std::async(std::launch::async,Y(),1.2);  
auto f7=std::async(std::launch::deferred,baz,std::ref(x));  
auto f8=std::async(  
    std::launch::deferred | std::launch::async,  
    baz,std::ref(x));  
auto f9=std::async(baz,std::ref(x));  
f7.wait();
```

← **Run in new thread**

← **Run in wait() or get()**

← **Implementation chooses**

← **Invoke deferred function**

В некоторых случаях, тем не менее, необходимо обеспечить, чтобы асинхронная операция обязательно выполнялась в рамках отдельного потока. Тогда используется флаг `std::launch::async`.

По умолчанию



# std::packaged\_task

```
template< class R, class ...Args >  
class packaged_task<R(Args...)>;
```

Связывает сущность типа Callable с аргументами, которая может быть потом вызвана посредством оператора ():

```
void operator()( ArgTypes... args );
```

Результат типа R возвращается через фьючерс с помощью метода

```
std::future<R> get_future();
```

# Задания `packaged_task`

- Могут передаваться между потоками, тем самым облегчая создание различных схем планирования заданий
- **`std::packaged_task<T(...)>`** связывает вызываемый объект (Callable) с **`std::future<T>`**
- Шаблонный параметр представляет собой сигнатуру функции, из которой выводятся тип результата

# Пример специализации packaged\_task

```
template<>
class packaged_task<std::string(std::vector<char>*,int)>
{
public:
    template<typename Callable>
    explicit packaged_task(Callable&& f);
    std::future<std::string> get_future();
    void operator()(std::vector<char>*,int);
};
```

# Пример использования packaged\_task

```
#include <deque>
#include <mutex>
#include <future>
#include <thread>
#include <utility>
std::mutex m;
std::deque<std::packaged_task<void()> > tasks;
bool gui_shutdown_message_received();
void get_and_process_gui_message();
```

# Пример использования packaged\_task

```
void gui_thread()
{
    while(!gui_shutdown_message_received())
    {
        get_and_process_gui_message();
        std::packaged_task<void()> task;
        {
            std::lock_guard<std::mutex> lk(m);
            if(tasks.empty())
                continue;
            task=std::move(tasks.front());
            tasks.pop_front();
        }
        task();
    }
}
```

# Пример использования packaged\_task

```
template<typename Func>
std::future<void> post_task_for_gui_thread(Func f)
{
    std::packaged_task<void()> task(f);
    std::future<void> res=task.get_future();
    std::lock_guard<std::mutex> lk(m);
    tasks.push_back(std::move(task));
    return res;
}
```

# std::promise

Позволяет явным образом сохранять в разделяемый объект (shared state) значение или исключение, которое в дальнейшем может быть получено с помощью std::future.

# Конструкторы `std::promise`

- Конструкторы
  - `promise()`; создает объект с пустым разделяемым состоянием;
  - `template< class Alloc > promise( std::allocator_arg_t, const Alloc& alloc );` создает объект с помощью заданного аллокатора;
  - `promise( promise&& other );` - конструктор перемещения;
  - `promise( const promise& other ) = delete;` конструктор копирования запрещен.



# Методы std::promise

- **std::future<T> get\_future();** - возвращает фьючерс, соответствующий изменяемому состоянию
- установка значения:  
**void set\_value( const R& value );**  
**void set\_value( R&& value );**  
**void set\_value( R& value );**      - для promise<R>  
**void set\_value();**                      - для promise<void>  
может вызываться только один раз (при повторном вызове возникает исключение)
- **установка значения по выходу из потока:**  
**void set\_value\_at\_thread\_exit( const R& value );**  
**void set\_value\_at\_thread\_exit( R&& value );**  
**void set\_value\_at\_thread\_exit( R& value );**  
**void set\_value\_at\_thread\_exit();**

# Пример использования std::promise

```
#include <vector>
#include <thread>
#include <future>
#include <numeric>
#include <iostream>
#include <chrono>

void accumulate(std::vector<int>::iterator first,
               std::vector<int>::iterator last,
               std::promise<int> accumulate_promise)
{
    int sum = std::accumulate(first, last, 0);
    accumulate_promise.set_value(sum); // Notify future
}
```

# Пример использования std::promise

```
int main()
{
    std::vector<int> numbers = { 1, 2, 3, 4, 5, 6 };
    std::promise<int> accumulate_promise;
    std::future<int> accumulate_future = accumulate_promise.get_future();
    std::thread work_thread(accumulate, numbers.begin(), numbers.end(),
                           std::move(accumulate_promise));
    accumulate_future.wait();
    std::cout << "result=" << accumulate_future.get() << '\n';
    work_thread.join(); }
```

# Второй пример использования `std::promise::set_value`

```
#include <iostream>
#include <future>
#include <thread>
#include <chrono>

int main()
{
    std::promise<int> p;
    std::future<int> f = p.get_future();
    std::chrono::time_point<std::chrono::system_clock> start, end;
    start = std::chrono::system_clock::now();
    std::thread([&p] {
        p.set_value(9);
        std::this_thread::sleep_for(std::chrono::seconds(3));
    }).detach();
```

# Второй пример использования std::promise::set\_value

```
std::cout << "Waiting..." << std::flush;
std::cout << "Done!\nResult is: " << f.get() << '\n';
end = std::chrono::system_clock::now();
std::cout << "Elapsed " << std::chrono::duration_cast<std::chrono::seconds>
(end-start).count() << " seconds\n";
}
```

```
posypkin@mikhail:~/exp/sandbox/c11/promise$ ./promise1
Waiting...Done!
Result is: 9
Elapsed 0 seconds
```

# Пример использования set\_value\_at\_thread\_exit

```
#include <iostream>
#include <future>
#include <thread>
#include <chrono>

int main()
{
    std::promise<int> p;
    std::future<int> f = p.get_future();
    std::chrono::time_point<std::chrono::system_clock> start, end;
    start = std::chrono::system_clock::now();
    std::thread([&p] {
        p.set_value_at_thread_exit(9);
        std::this_thread::sleep_for(std::chrono::seconds(3));
    }).detach();
```

# Пример использования set\_value\_at\_thread\_exit

```
std::cout << "Waiting..." << std::flush;  
std::cout << "Done!\nResult is: " << f.get() << '\n';  
end = std::chrono::system_clock::now();  
std::cout << "Elapsed " << std::chrono::duration_cast<std::chrono::seconds>  
(end-start).count() << " seconds\n";  
}
```

```
posypkin@mikhail:~/exp/sandbox/c11/promise$ ./promise1  
Waiting...Done!  
Result is: 9  
Elapsed 3 seconds
```

**СПАСИБО ЗА ВНИМАНИЕ!**  
**ВОПРОСЫ?**





# Умножение матриц

```
#include <stdlib.h>
#include <iostream>
#include <thread>
#include <vector>
#include <chrono>

// Number of matrix rows and
// columns
int n;

// Number of threads
int nt;

double *a, *b, *c;

// Thread-firendly matrix multiply
// tn - thread number
void mmul(int tn) {
    for(int i = tn; i < n; i += nt) {
        for(int j = 0; j < n; j++) {
            double v = 0;
            for(int k = 0; k < n; k++)
                v += a[i*n + k] * b[k*n + j];
            c[i*n + j] = v;
        }
    }
}
```



```
// Fill in matrix
void init_matr()
{
    a = new double[n * n];
    b = new double[n * n];
    c = new double[n * n];
    for(int i = 0; i < n; i++) {
        for(int j = 0; j < n; j++) {
            a[i*n + j] = 1;
            b[i*n + j] = 2;
        }
    }
}
```

```
// Check result
void check() {
    int check_step = n / 8;
    for(int i = 0; i < n; i +=
check_step) {
        for(int j = 0; j < n; j +=
check_step) {
            if(c[i*n + j] != 2 * n) {
                std::cerr << "Error
detected\n";
                exit(-1);
            }
        }
    }
}
```

```

main(int argc, char* argv[]) {
    if(argc != 3) {
        std::cerr << "usage: mmul matrix_dim
num_of_thread\n";
        exit(-1);
    }
    n = atoi(argv[1]);
    nt = atoi(argv[2]);
    init_matr();
    std::vector< std::thread > thv;

```

```

std::chrono::time_point<std::chrono::sy
stem_clock> start, end;
    start =
std::chrono::system_clock::now();
    for(int i = 0; i < nt; i++) {
        std::thread th(mmul, i);
        thv.push_back(std::move(th));
    }
    for(auto& t : thv) {
        t.join();
    }
    end =
std::chrono::system_clock::now();
    std::chrono::duration<double> dur =
end - start;
    std::cout << "Matrix multiply took " <<
dur.count() << "\n";
    check();
}

```