Parallel computing

Pthread

Objetivos

- Crear threads
- Sincronizar threads
- Acceder a secciones críticas
- Cache
- Usar Posix threads

Pthread

- Posix Thread Libreria multithreading
 - Sistemas Unix
- STD::Thread c++11
- Windows thread

Compilar

\$ gcc -g -Wall -o hello hello.c -lpthread

```
#include <stdio.h>
   #include <stdlib.h>
   #include <pthread.h>
   /* Global variable: accessible to all threads */
   int thread_count:
   void* Hello(void* rank): /* Thread function */
9
   int main(int argc, char* argv[]) {
11
                  thread; /* Use long in case of a 64-bit system */
      pthread_t* thread_handles;
12
13
      /* Get number of threads from command line */
14
15
      thread_count = strtol(argv[1], NULL, 10):
16
17
      thread_handles = malloc (thread_count*sizeof(pthread_t));
18
```

```
19
      for (thread = 0: thread < thread_count: thread++)
20
          pthread_create(&thread_handles[thread], NULL.
              Hello. (void*) thread):
21
22
23
       printf("Hello from the main thread\n");
24
25
      for (thread = 0; thread < thread_count; thread++)</pre>
          pthread_join(thread_handles[thread], NULL):
26
27
      free(thread_handles):
       return 0:
      /* main */
31
   void* Hello(void* rank) {
       long my_rank = (long) rank
33
             /* Use long in case of 64-bit system */
34
       printf("Hello from thread %ld of %d\n", my_rank,
35
             thread_count):
36
37
       return NULL:
       /* Hello */
```

Ejecutar

- \$./hello 4
- Hello from the main thread
- Hello from thread 0 of 4
- Hello from thread 1 of 4
- Hello from thread 2 of 4
- Hello from thread 3 of 4

Codigo

- Variables globales (warning)
 - Ej. thread_count es compartida entre todos los threads
- Cada thread tiene su propio stack

Crear threads

- Objeto pthread_t por cada thread
 - Contiene informacion especifica de cada thread
 - Opaco (especifico al sistema)

- pthread_t tiene que ser asignado con malloc antes de llamar esta funcion
- El tercer parámetro es la función que tendrá que ejecutar el thread
- El cuarto parámetro corresponde al parámetro que queremos mandar a la rutina

Crear threads

- Las funciones start_routine tienen que tener el siguiente prototipo
- void* thread_function(void* args_p);
- void* cast cualquier tipo de puntero
 - struct para pasar varios datos
- Podemos siempre pasar el rank de cada thread (util para debug)

Crear threads

- El ejemplo anterior, los threads ejecutan todos el mismo código.
- Cada thread podría ejecutar una función diferente.

Ejecución de los threads

- El thread que ejecuta el main
 - o main thread
- Mientras los threads creados con pthread_create ejecutan la función hello
 - Los threads retornan NULL
- No se especifica dónde y cuando se ejecuta el thread
- sobre un sistema cargado todos los threads podrían ejecutarse sobre el mismo núcleo
- En cambio sí un núcleo está libre el sistema lo aprovechara

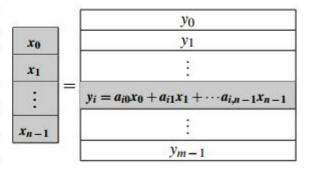
Parar los threads

Tenemos que llamar un pthread_join por cada thread creado

```
int pthread_join(
    pthread_t thread /* in */,
    void** ret_val_p /* out */);
```

Ej. Multiplicación Matriz - Vector

a_{00}	<i>a</i> ₀₁		$a_{0,n-1}$
a_{10}	<i>a</i> ₁₁		$a_{1,n-1}$
:	:		:
a_{i0}	a _{i1}		$a_{i,n-1}$
÷	:		:
$a_{m-1,0}$	$a_{m-1,1}$	•••	$a_{m-1,n-1}$



```
/* For each row of A */
for (i = 0; i < m; i++) {
   y[i] = 0.0;
   /* For each element of the row and each element of x */
   for (j = 0; j < n; j++)
      y[i] += A[i][j]* x[j];
}</pre>
```

Ej. Multiplicación Matriz - Vecctor

Repartir el trabajo:

 \circ ej: 3 threads, n = 6

Thread	Components of y			
0	y[0],	y[1]		
1	y[2],			
2	y[4].	y[5]		

Cada thread ejecutara para todos sua "i" atribuidos :

```
y[i] = 0.0;
for (j = 0; j < n; j++)
  y[i] += A[i][j]*x[j];</pre>
```

Ej. Multiplicación Matriz - Vecctor

- y, x, A son compartidos
- calcular los "i" de cada thread "q"
- inicio = q * m/t
 - m = tamaño del vector resultante
 - t cuantos vectores
- fin = ((q+1)*m/t)-1

```
y[i] = 0.0;

for (j = 0; j < n; j++)

y[i] += A[i][j]*x[j];
```

```
void* Pth_mat_vect(void* rank) {
  long my_rank = (long) rank;
  int i, j;
  int local_m = m/thread_count;
  int my_first_row = my_rank*local_m;
  int my_last_row = (my_rank+1)*local_m - 1;

for (i = my_first_row; i <= my_last_row; i++) {
    y[i] = 0.0;
    for (j = 0; j < n; j++)
        y[i] += A[i][j]*x[j];
}

return NULL;
} /* Pth_mat_vect */</pre>
```

Secciones críticas

- En el ejemplo anterior no existe ninguna
 - race condition
- Como en OpenMP podemos estimar Pi

```
double factor = 1.0;
double sum = 0.0;
for (i = 0; i < n; i++, factor = -factor) {
    sum += factor/(2*i+1);
}
pi = 4.0*sum;</pre>
```

Secciones críticas

Como en OpenMP podemos estimar Pi

```
void* Thread_sum(void* rank) {
       long my_rank = (long) rank;
       double factor:
       long long i;
       long long my_n = n/thread_count;
       long long my_first_i = my_n*my_rank;
       long long my_last_i = my_first_i + my_n;
9
       if (my_first_i % 2 == 0) /* my_first_i is even */
10
          factor = 1.0:
11
       else /* my_first_i is odd */
          factor = -1.0:
12
13
14
       for (i = my_first_i; i < my_last_i; i++, factor = -factor) {</pre>
          sum += factor/(2*i+1):
15
16
17
       return NULL:
18
       /* Thread_sum */
```

	п				
	105	10 ⁶	107	108	
π	3.14159	3.141593	3.1415927	3.14159265	
1 Thread	3.14158	3.141592	3.1415926	3.14159264	
2 Threads	3.14158	3.141480	3.1413692	3.14164686	

Busy waiting

sección crítica

```
    x=x+y
    y = Compute(my rank);
    while (flag != my rank);
    x = x + y;
    flag++;
```

Busy waiting

- Esta solución corrige los problemas
 - o con n= 10^8 el tiempo p = 7^* t serial

```
void* Thread_sum(void* rank) {
       long my_rank = (long) rank;
      double factor:
      long long i;
      long long my_n = n/thread_count;
      long long my_first_i = my_n*my_rank;
      long long my_last_i = my_first_i + my_n;
      if (my_first_i \% 2 == 0)
          factor = 1.0:
11
      else
          factor = -1.0:
12
13
      for (i = my_first_i; i < my_last_i; i++, factor = -factor) {</pre>
14
          while (flag != my_rank);
15
16
          sum += factor/(2*i+1):
          flag = (flag+1) % thread_count;
17
18
19
20
       return NULL:
      /* Thread_sum */
```

- Mutual exclusion
- Excluye los otros threads de una zona
- pthread_mutex_t
- int pthread_mutex_init(pthread mutex_t* mutex_p, NULL);
- int pthread_mutex_destroy(pthread_mutex_t* mutex_p);

- Acceder a una zona crítica (lock)
 - o int pthread_mutex_lock(pthread_mutex_t* mutex_p);
- Dejar una zona crítica (unlock)
 - int pthread_mutex_unlock(pthread_mutex_t* mutex_p);
- lock
 - o causa la **espera** de los demas threads
 - El sistema decide quién será el siguiente thread en obtener el acceso
 - No hay un orden predefinido
- unlock
 - o notifica el sistema que la sección crítica está libre

```
void* Thread_sum(void* rank) {
 2
       long my_rank = (long) rank;
       double factor:
       long long i;
       long long my_n = n/thread_count;
       long long my_first_i = my_n*my_rank;
       long long my_last_i = my_first_i + my_n;
       double my_sum = 0.0;
9
10
       if (my_first_i % 2 == 0)
11
          factor = 1.0:
       else
12
13
          factor = -1.0:
14
15
       for (i = my_first_i; i < my_last_i; i++, factor = -factor) {</pre>
16
          my_sum += factor/(2*i+1);
17
       pthread_mutex_lock(&mutex):
18
19
       sum += my_sum:
       pthread_mutex_unlock(&mutex);
20
21
22
       return NULL:
23
       /* Thread_sum */
```

Cuando la cantidad de thread <=8 (cantidad de cores)

$$rac{T_{
m serial}}{T_{
m parallel}} pprox {
m thread_count}$$

0	Eficiencia quasi idéal	

- Cuando la cantidad de threads > cores
 - o Busy waiting empeora, ej: 2 cores 5 threads (flag !!)

	flag	Thread				
Time		0	1	2	3	4
0	0	crit sect	busy-wait	susp	susp	susp
1	1	terminate	crit sect	susp	busy-wait	susp
2	2	-	terminate	susp	busy-wait	busy-wait
:					:	:
?	2	-	-	crit sect	susp	busy-wait

Threads	Busy-Wait	Mutex
1	2.90	2.90
2	1.45	1.45
4	0.73	0.73
8	0.38	0.38
16	0.50	0.38
32	0.80	0.40
64	3.56	0.38
		22

- Cómo obligar los thread a ejecutarse en una secuencia?
 - Ej.: Operaciones no conmutativas como la multiplicación matricial

```
/* n and product_matrix are shared and initialized by the main
thread */
/* product_matrix is initialized to be the identity matrix */
void* Thread_work(void* rank) {
   long my_rank = (long) rank;
   matrix_t my_mat = Allocate_matrix(n);
   Generate_matrix(my_mat);
   pthread_mutex_lock(&mutex);
   Multiply_matrix(product_mat, my_mat);
   pthread_mutex_unlock(&mutex);
   Free_matrix(&my_mat);
   return NULL;
} /* Thread_work */
```

- Cómo obligar los thread a ejecutarse en una secuencia?
 - Ej.: Cada thread tiene que mandar un mensaje a otro thread
 - Thread 0 -> thread 1 -> thread 2 -> thread t -> thread 0
 - Un thread después de recibir su mensaje puede terminar
 - Array "messages" char** de tamaño t
 - Un thread agrega un mensaje messages[dest]=myMessage
 - Después lee el mensaje que recibió

- Cómo obligar los thread a ejecutarse en una secuencia?
 - Ej.: Cada thread tiene que mandar un mensaje a otro thread

```
/* messages has type char**. It's allocated in main. */
   /* Each entry is set to NULL in main.
   void* Send_msg(void* rank) {
       long my_rank = (long) rank:
      long dest = (my_rank + 1) % thread_count;
      long source = (my_rank + thread_count - 1) % thread_count;
      char* my_msg = malloc(MSG_MAX*sizeof(char));
      sprintf(my_msg, "Hello to %ld from %ld", dest, my_rank);
10
      messages[dest] = my_msg;
11
      if (messages[my_rank] != NULL)
12
         printf("Thread %ld > %s\n", my_rank, messages[my_rank]);
13
14
      else
          printf("Thread %ld > No message from %ld\n", my_rank,
15
                source):
16
17
      return NULL:
       /* Send_msq */
```

- Cómo obligar los thread a ejecutarse en una secuencia?
 - Ej.: Cada thread tiene que mandar un mensaje a otro thread
 - Es muy probable que el Thread 0 intenté leer su mensaje antes que el último thread le haya escrito

- Ej.: Cada thread tiene que mandar un mensaje a otro thread
- Podemos hacer esto con mutex ?

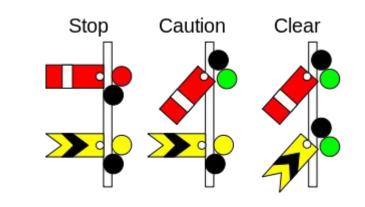
```
pthread_mutex_lock(mutex[dest]);

messages[dest] = my_msg;
pthread_mutex_unlock(mutex[dest]);

pthread_mutex_lock(mutex[my_rank]);
printf("Thread %ld > %s\n", my_rank, messages[my_rank]);
```

Thread 0 podría llegar a la linea 7 antes que thread t llegue a la linea 2

- Solución: semaphores (Dijkstra)
 - Carriles de tren
 - Unsigned int
 - Muchas veces solo 0-1 (binary semaphore)
 - 0 locked 1 unlocked
 - sem_wait bloquea si sempahore = 0 sino decrementa el semaphore
 - sem_post para incrementar el semaphore
 - o los semphores no tienen "dueños" entonces el main thread puede inicializar todos los semphores a 0



Solución: semaphores

```
1 /* messages is allocated and initialized to NULL in main */
   /* semaphores is allocated and initialized to 0 (locked) in
         main */
   void* Send_msg(void* rank) {
      long my_rank = (long) rank;
      long dest = (my_rank + 1) % thread_count;
      char* my_msg = malloc(MSG_MAX*sizeof(char));
      sprintf(my_msg, "Hello to %ld from %ld", dest, my_rank);
9
      messages[dest] = my_msg;
      sem_post(&semaphores[dest])
10
            /* ''Unlock'' the semaphore of dest */
11
12
      /* Wait for our semaphore to be unlocked */
13
      sem_wait(&semaphores[my_rank]);
      printf("Thread %ld > %s\n", my_rank, messages[my_rank]);
14
15
16
      return NULL:
      /* Send_msq */
```

- Solución: semaphores
- #include <semaphore.h>

```
int sem_init(
    sem_t* semaphore_p /* out */,
    int shared /* in */,
    unsigned initial_val /* in */);

int sem_destroy(sem_t* semaphore_p /* in/out */);
int sem_post(sem_t* semaphore_p /* in/out */);
int sem_wait(sem_t* semaphore_p /* in/out */);
```

- Sincronizar los threads
 - Asegurar que todos lleguen al mismo lugar en el programa
 - barrera
 - Ej. medir el tiempo del thread mas lento

```
/* Shared */
double elapsed_time;
...
/* Private */
double my_start, my_finish, my_elapsed;
...
Synchronize threads;
Store current time in my_start;
/* Execute timed code */
...
Store current time in my_finish;
my_elapsed = my_finish - my_start;
elapsed = Maximum of my_elapsed values;
```

- Sincronizar los threads
 - Asegurar que todos lleguen al mismo lugar en el programa
 - barrera
 - Ej. depuración

```
point in program we want to reach;
barrier;
if (my_rank == 0) {
   printf("All threads reached this point\n");
   fflush(stdout);
}
```

- Según los sistemas no disponen de alguna función de barrera
 - Hay que implementarla
 - busy-waiting o mutex
 - semaphore
 - variable de condition

- Busy waiting y mutex
 - Contador de threads que pasan por una zona critica

```
/* Shared and initialized by the main thread */
int counter; /* Initialize to 0 */
int thread_count;
pthread_mutex_t barrier_mutex;
...

void* Thread_work(. . .) {
    ...
    /* Barrier */
    pthread_mutex_lock(&barrier_mutex);
    counter++;
    pthread_mutex_unlock(&barrier_mutex);
    while (counter < thread_count);
    ...
}</pre>
```

- Se gasta mucho CPU en el while aun + si threads > cores
- Que pasa si más adelante se necesita otra barrera
 - o no es posible resetear counter (o al menos muy peligroso)
 - Los threads se podrían quedar en el while

Semaphore

```
sem_t count_sem: /* Initialize to 1 */
sem_t barrier_sem; /* Initialize to 0 */
. . .
void* Thread_work(...) {
   /* Barrier */
   sem_wait(&count_sem):
   if (counter == thread_count-1) {
      counter = 0:
      sem_post(&count_sem);
      for (j = 0; j < thread_count-1; j++)
         sem_post(&barrier_sem);
   } else {
      counter++:
      sem_post(&count_sem);
      sem_wait(&barrier_sem);
```

- 2 Semaphores
 - count_sem: protege counter 1
 - barrier_sem: impide los threads de terminar 0
- Cuando un thread entra en el "else"
 - incrementa el contador libera count_sem
 - se queda a la espera de su liberación barrier_sem
- Cuando el ultimo thread entra
 - resetea el contador
 - libera el count_sem
 - libera todo los threads uno por uno

```
sem_t count_sem:
                   /* Initialize to 1 */
sem_t barrier_sem: /* Initialize to 0 */
void* Thread_work(...) {
   /* Barrier */
   sem_wait(&count_sem):
   if (counter == thread_count-1) {
      counter = 0:
      sem_post(&count_sem);
      for (j = 0; j < thread_count-1; j++)
         sem_post(&barrier_sem);
   } else {
      counter++:
      sem_post(&count_sem);
      sem_wait(&barrier_sem):
```

- el for libera todos los threads llamando a sem_post
 - sem_post incrementa el unside int
 - no importa si se llama muchas sem_post sin sem_wait
 - sem_wait decrementa hasta 0
- No ocupa el CPU por gusto
- Podemos reutilizar la barrera?
- reinicializamos el contador OK
- count_sem a 1 OK
- barrier_sem No estamos seguros

```
sem_t count_sem:
                    /* Initialize to 1 */
sem_t barrier_sem: /* Initialize to 0 */
void* Thread_work(...) {
   /* Barrier */
   sem_wait(&count_sem):
   if (counter == thread_count-1) {
      counter = 0:
      sem_post(&count_sem):
      for (j = 0; j < thread_count-1; j++)
         sem_post(&barrier_sem);
   } else {
      counter++:
      sem_post(&count_sem);
      sem_wait(&barrier_sem):
```

- Variable condición pthread_cond_t
- Permite la suspensión de un thread hasta que un evento o condición ocurra
- Asocia a un mutex
- Desbloquear un thread bloqueado
 - o int pthread_cond_signal(pthread_cond_t* cond_var_p);
- Desbloquea todo los threads
 - o int pthread_cond_broadcast(pthread_cond_t* cond_var_p);
- Desbloquea un mutex
 - int pthread_cond_wait(pthread_cond_t* cond_var_p, pthread_mutex_t* mutex_p);
 - Esto causa que el thread ejecutando se bloquee hasta ser desbloqueado por otro thread (boradcast o signal)

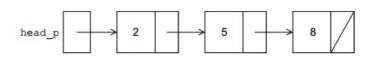
```
pthread_mutex_unlock(&mutex_p);
wait_on_signal(&cond_var_p);
pthread_mutex_lock(&mutex_p);
```

- Un thread entra en el else y llama a cond_wait
 - el thread se queda esperando y se desbloque el mutex
 - o asi para todos menos el ultimo
- El último thread entra en el if
 - o reinicializa el contador
 - Desbloquea todos los threads un por uno

```
/* Shared */
int counter = 0:
pthread_mutex_t mutex;
pthread_cond_t cond_var;
void* Thread_work(. . .) {
   /* Barrier */
   pthread_mutex_lock(&mutex);
   counter++:
   if (counter == thread_count) {
      counter = 0:
      pthread_cond_broadcast(&cond_var);
   } else {
      pthread_cond_wait(&cond_var, &mutex)
   pthread_mutex_unlock(&mutex);
```

- Compartir una estructura de datos entre threads
- Ej. lista enlazada
 - o member, insert, delete

```
struct list_node_s {
   int data;
   struct list_node_s* next;
}
```



Member

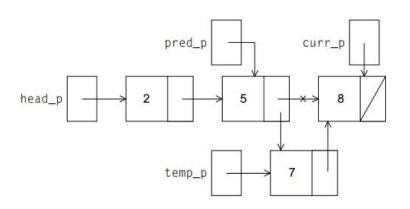
```
int Member(int value, struct list_node_s* head_p) {
   struct list_node_s* curr_p = head_p;

while (curr_p != NULL && curr_p->data < value)
   curr_p = curr_p->next;

if (curr_p == NULL || curr_p->data > value) {
   return 0;
   } else {
   return 1;
}

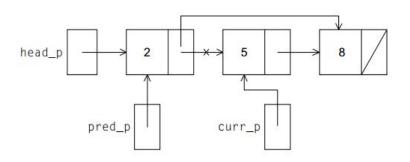
/* Member */
```

Insert



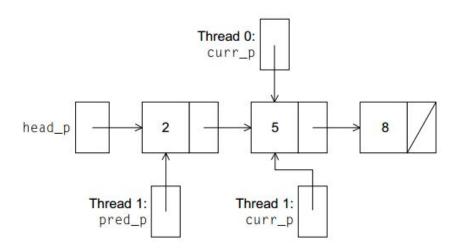
```
int Insert(int value, struct list_node_s** head_p) {
      struct list_node_s* curr_p = *head_p;
      struct list_node_s* pred_p = NULL;
      struct list_node_s* temp_p;
      while (curr_p != NULL && curr_p->data < value) {
          pred_p = curr_p;
          curr_p = curr_p->next;
9
10
11
      if (curr_p == NULL || curr_p->data > value) {
12
          temp_p = malloc(sizeof(struct list_node_s));
13
          temp_p->data = value;
14
          temp_p->next = curr_p;
15
         if (pred_p == NULL) /* New first node */
16
             *head_p = temp_p;
         else
18
             pred_p->next = temp_p;
19
          return 1:
       } else { /* Value already in list */
20
21
          return 0:
22
23
       /* Insert */
```

delete



```
int Delete(int value, struct list_node_s** head_p) {
      struct list_node_s* curr_p = *head_p;
      struct list_node_s* pred_p = NULL;
      while (curr_p != NULL && curr_p->data < value) {
          pred_p = curr_p;
         curr_p = curr_p->next;
10
      if (curr_p != NULL && curr_p->data == value) {
         if (pred_p == NULL) { /* Deleting first node in list */
11
             *head_p = curr_p->next;
13
            free(curr_p);
14
         } else {
15
             pred_p->next = curr_p->next;
16
            free(curr_p);
17
18
         return 1:
19
      } else { /* Value isn't in list */
20
         return 0:
21
       /* Delete */
```

- Que pasa cuando usamos esta estructura con varios threads
- Ej. Thread 0 member(5), thread 1 delete(5)



- Que pasa cuando usamos esta estructura con varios threads
- Ej. Thread 0 member(5), thread 1 delete(5)
 - thread 0 retorne True mientras no existe el 5
 - o thread 0 quiera acceder al 5 mientras la memoria ha sido liberada (crash)

Solución rapida

```
Pthread_mutex_lock(&list_mutex);
Member(value);
Pthread_mutex_unlock(&list_mutex);
```

- Solución 1: los threads acceden a la lista uno a la vez
- Solución 2 : los threads acceden a un nodo uno a la vez
- Alternativa : read-write-lock
 - Dos funciones de lock
 - read
 - write
 - Múltiples threads pueden obtener el lock para el read
 - Uno solamente puede obtener el lock para write
 - Si un thread ha obtenido el look para el read entonces ninguno obtendrá un lock para el writing
 - Si un thread ha obtenido el look para el write entonces ninguno obtendrá un lock para el writing/reading

read-write-lock

- int pthread_rwlock_init(pthread rwlock t* rwlock p,NULL);
- int pthread_rwlock_destroy(pthread rwlock t* rwlock p);
- o int pthread_rwlock_rdlock(pthread_rwlock_t* rwlock_p);
- int pthread_rwlock_wrlock(pthread_rwlock_t* rwlock_p);
- o int pthread_rwlock_unlock(pthread_rwlock_t* rwlock_p);

```
pthread_rwlock_rdlock(&rwlock);
Member(value);
pthread_rwlock_unlock(&rwlock);
...
pthread_rwlock_wrlock(&rwlock);
Insert(value);
pthread_rwlock_unlock(&rwlock);
...
pthread_rwlock_wrlock(&rwlock);
Delete(value);
pthread_rwlock_unlock(&rwlock);
```

Performance

- o list size 1000
- o 100 000 ops
- o 99.9% member 0.05% insert 0.05% delete

	Number of Threads				
Implementation	1	2	4	8	
Read-Write Locks	0.213	0.123	0.098	0.115	
One Mutex for Entire List	0.211	0.450	0.385	0.457	
One Mutex per Node	1.680	5.700	3.450	2.700	

Performance

- o list size 1000
- o 100 000 ops
- o 80% member 10% insert 10% delete

Implementation	Number of Threads			
	1	2	4	8
Read-Write Locks	2.48	4.97	4.69	4.71
One Mutex for Entire List	2.50	5.13	5.04	5.11
One Mutex per Node	12.00	29.60	17.00	12.00

thread c++11

Crear thread std::thread t1(call_from_thread); Join t1.join() void call_from_thread(int tid) { std::cout << "Launched by thread " << tid << std::endl; int main() { std::thread t[num_threads]; for (int i = 0; $i < num_threads$; ++i) { t[i] = std::thread(call_from_thread, i);