Document Number: N4104
Date: 2014-07-04
Revises: N3960

Editor: Jared Hoberock

NVIDIA Corporation jhoberock@nvidia.com

Working Draft, Technical Specification for C++ Extensions for Parallelism

Note: this is an early draft. It's known to be incomplet and incorrekt, and it has lots of bad formatting.

Contents

1	Gene	eral
	1.1	Scope
	1.2	Normative references
	1.3	Namespaces and headers
		1.3.1 Terms and definitions
2	Exec	cution policies
	2.1	In general
	2.2	Header <experimental execution_policy=""> synopsis</experimental>
	2.3	Execution policy type trait
	2.4	Sequential execution policy
	2.5	Parallel execution policy
	2.6	Parallel+Vector execution policy
	2.7	Dynamic execution policy
		2.7.1 execution_policy construct/assign
		2.7.2 execution_policy object access
	2.8	Execution policy objects
3		llel exceptions
	3.1	Exception reporting behavior
_	3.2	Header <experimental exception_list=""> synopsis</experimental>
4		llel algorithms
	4.1	In general
		4.1.1 Requirements on user-provided function objects
		4.1.2 Effect of execution policies on algorithm execution
	4.0	4.1.3 ExecutionPolicy algorithm overloads
	4.2	Definitions
	4.3	Non-Numeric Parallel Algorithms
		4.3.1 Header <experimental algorithm=""> synopsis</experimental>
	4.4	4.3.2 For each
	4.4	Numeric Parallel Algorithms
		4.4.1 Header <experimental numeric=""> synopsis</experimental>
		4.4.2 Reduce
		4.4.4 Inclusive scan
		- 4.4.4 HIGHOLVC OUGH

1 General

[parallel.general]

1.1 Scope

[parallel.general.scope]

This Technical Specification describes requirements for implementations of an interface that computer programs written in the C++ programming language may use to invoke algorithms with parallel execution. The algorithms described by this Technical Specification are realizable across a broad class of computer architectures.

- ² This Technical Specification is non-normative. Some of the functionality described by this Technical Specification may be considered for standardization in a future version of C++, but it is not currently part of any C++ standard. Some of the functionality in this Technical Specification may never be standardized, and other functionality may be standardized in a substantially changed form.
- ³ The goal of this Technical Specification is to build widespread existing practice for parallelism in the C++ standard algorithms library. It gives advice on extensions to those vendors who wish to provide them.

1.2 Normative references

[parallel.general.references]

- ¹ The following referenced document is indispensable for the application of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.
 - ISO/IEC 14882: $-^1$, Programming Languages C++
- ² ISO/IEC 14882:— is herein called the *C++ Standard*. The library described in ISO/IEC 14882:— clauses 17-30 is herein called the *C++ Standard Library*. The C++ Standard Library components described in ISO/IEC 14882:— clauses 25, 26.7 and 20.7.2 are herein called the *C++ Standard Algorithms Library*.
- ³ Unless otherwise specified, the whole of the C++ Standard's Library introduction (C++14 §17) is included into this Technical Specification by reference.

1.3 Namespaces and headers

[parallel.general.namespaces]

- Since the extensions described in this Technical Specification are experimental and not part of the C++ Standard Library, they should not be declared directly within namespace std. Unless otherwise specified, all components described in this Technical Specification are declared in namespace std::experimental::parallel::v1.
 - [Note: Once standardized, the components described by this Technical Specification are expected to be promoted to namespace std. $end\ note$]
- ² Unless otherwise specified, references to such entities described in this Technical Specification are assumed to be qualified with std::experimental::parallel::v1, and references to entities described in the C++ Standard Library are assumed to be qualified with std::.
- Extensions that are expected to eventually be added to an existing header <meow> are provided inside the <experimental/meow> header, which shall include the standard contents of <meow> as if by

#include <meow>

1. To be published. Section references are relative to N3937.

§ 1.3

1.3.1 Terms and definitions

[parallel.general.defns]

¹ For the purposes of this document, the terms and definitions given in the C++ Standard and the following apply.

- ² A *parallel algorithm* is a function template described by this Technical Specification declared in namespace std::experimental::parallel::v1 with a formal template parameter named ExecutionPolicy.
- ³ Parallel algorithms access objects indirectly accessible via their arguments by invoking the following functions:
 - All operations of the categories of the iterators that the algorithm is instantiated with.
 - Functions on those sequence elements that are required by its specification.
 - User-provided function objects to be applied during the execution of the algorithm, if required by the specification.

These functions are herein called *element access functions*.

[*Example:* The sort function may invoke the following element access functions:

- Methods of the random-access iterator of the actual template argument, as per 24.2.7, as implied by the name of the template parameters RandomAccessIterator.
- The swap function on the elements of the sequence (as per 25.4.1.1 [sort]/2).
- The user-provided Compare function object.
- end example]

§ 1.3.1 4

2 Execution policies

[parallel.execpol]

2.1 In general

[parallel.execpol.general]

¹ This clause describes classes that are *execution policy* types. An object of an execution policy type indicates to an algorithm whether it is allowed to execute in parallel and expresses the requirements on the element access functions.

```
[ Example:
 std::vector<int> v = ...
 // standard sequential sort
 std::sort(vec.begin(), vec.end());
 using namespace std::experimental::parallel;
 // explicitly sequential sort
 sort(seq, v.begin(), v.end());
 // permitting parallel execution
 sort(par, v.begin(), v.end());
 // permitting vectorization as well
 sort(par vec, v.begin(), v.end());
 // sort with dynamically-selected execution
 size t threshold = ...
 execution policy exec = seq;
 if (v.size() > threshold)
 {
   exec = par;
 }
```

sort(exec, v.begin(), v.end());

— end example]

[*Note:* Because different parallel architectures may require idiosyncratic parameters for efficient execution, implementations of the Standard Library may provide additional execution policies to those described in this Technical Specification as extensions. — *end note*]

2.2 Header <experimental/execution_policy> synopsis [parallel.execpol.synopsis]

```
namespace std {
namespace experimental {
namespace parallel {
inline namespace v1 {
    // 2.3, Execution policy type trait
    template<class T> struct is_execution_policy;
    template<class T> constexpr bool is_execution_policy_v = is_execution_policy<T>::value;
```

§ 2.2 5

```
// 2.4, Sequential execution policy
class sequential_execution_policy;

// 2.5, Parallel execution policy
class parallel_execution_policy;

// 2.6, Parallel+Vector execution policy
class parallel_vector_execution_policy;

// 2.7, Dynamic execution policy
class execution_policy;
}
}
}
```

2.3 Execution policy type trait

[parallel.execpol.type]

template<class T> struct is_execution_policy { see below };

- ¹ is_execution_policy can be used to detect parallel execution policies for the purpose of excluding function signatures from otherwise ambiguous overload resolution participation.
- ² is_execution_policy<T> shall be a UnaryTypeTrait with a BaseCharacteristic of true_type if T is the type of a standard or implementation-defined execution policy, otherwise false_type.
- ³ The behavior of a program that adds specializations for is execution policy is undefined.

2.4 Sequential execution policy

[parallel.execpol.seq]

```
class sequential execution policy{ unspecified };
```

¹ The class sequential_execution_policy is an execution policy type used as a unique type to disambiguate parallel algorithm overloading and require that a parallel algorithm's execution may not be parallelized.

2.5 Parallel execution policy

[parallel.execpol.par]

```
class parallel execution policy{ unspecified };
```

¹ The class parallel_execution_policy is an execution policy type used as a unique type to disambiguate parallel algorithm overloading and indicate that a parallel algorithm's execution may be parallelized.

2.6 Parallel+Vector execution policy

[parallel.execpol.vec]

```
class parallel vector execution policy{ unspecified };
```

¹ The class class parallel_vector_execution_policy is an execution policy type used as a unique type to disambiguate parallel algorithm overloading and indicate that a parallel algorithm's execution may be vectorized and parallelized.

§ 2.6

2.7 Dynamic execution policy

[parallel.execpol.dynamic]

```
class execution_policy
{
  public:
    // 2.7.1, execution_policy construct/assign
    template<class T> execution_policy(const T& exec);
    template<class T> execution_policy& operator=(const T& exec);

// 2.7.2, execution_policy object access
  template<class T> T* get() noexcept;
  template<class T> const T* get() const noexcept;
};
```

¹ The class execution_policy is a container for execution policy objects. execution_policy allows dynamic control over standard algorithm execution.

[Example:

```
std::vector<float> sort_me = ...

using namespace std::experimental::parallel;
execution_policy exec = seq;

if(sort_me.size() > threshold)
{
   exec = std::par;
}

std::sort(exec, std::begin(sort_me), std::end(sort_me));
— end example]
```

Objects of type execution_policy shall be constructible and assignable from objects of type T for which is execution policy<T>::value is true.

2.7.1 execution policy construct/assign

[parallel.execpol.con]

- 1 template<class T> execution_policy(const T& exec);
 - ² Effects: Constructs an execution policy object with a copy of exec's state.
 - Remarks: This constructor shall not participate in overload resolution unless is_execution_policy<T>::value is true.
- 4 template<class T> execution policy& operator=(const T& exec);
 - ⁵ Effects: Assigns a copy of exec's state to *this.
 - 6 Returns: *this.

§ 2.7.1 7

2.7.2 execution_policy object access

[parallel.execpol.access]

```
const type_info& type() const noexcept;
```

² *Returns:* typeid(T), such that T is the type of the execution policy object contained by *this.

```
3 template<class T> T* get() noexcept;
    template<class T> const T* get() const noexcept;
```

- 4 Returns: If target_type() == typeid(T), a pointer to the stored execution policy object; otherwise a null pointer.
- ⁵ Requires: is_execution_policy<T>::value iS true.

2.8 Execution policy objects

[parallel.execpol.objects]

```
constexpr sequential_execution_policy seq{};
constexpr parallel_execution_policy par{};
constexpr parallel_vector_execution_policy par_vec{};
```

¹ The header <experimental/execution_policy> declares a global object associated with each type of execution policy defined by this Technical Specification.

§ 2.8

3 Parallel exceptions

[parallel.exceptions]

3.1 Exception reporting behavior

[parallel.exceptions.behavior]

¹ During the execution of a standard parallel algorithm, if temporary memory resources are required and none are available, the algorithm throws a std::bad_alloc exception.

- ² During the execution of a standard parallel algorithm, if the invocation of an element access function terminates with an uncaught exception, the behavior of the program is determined by the type of execution policy used to invoke the algorithm:
 - If the execution policy object is of type class parallel_vector_execution_policy, std::terminate shall be called.
 - If the execution policy object is of type sequential_execution_policy or parallel_execution_policy, the execution of the algorithm terminates with an exception_list exception. All uncaught exceptions thrown during the invocations of element access functions shall be contained in the exception_list.

[*Note:* For example, the number of invocations of the user-provided function object in for_each is unspecified. When for_each is executed sequentially, only one exception will be contained in the exception_list object. — *end note*]

[Note: These guarantees imply that, unless the algorithm has failed to allocate memory and terminated with std::bad_alloc, all exceptions thrown during the execution of the algorithm are communicated to the caller. It is unspecified whether an algorithm implementation will "forge ahead" after encountering and capturing a user exception.

— end note]

[Note: The algorithm may terminate with the std::bad_alloc exception even if one or more user-provided function objects have terminated with an exception. For example, this can happen when an algorithm fails to allocate memory while creating or adding elements to the exception_list object. — end note]

 If the execution policy object is of any other type, the behavior is implementationdefined.

3.2 Header <experimental/exception list> synopsis

3

[parallel.exceptions.synopsis]

```
namespace std {
namespace experimental {
namespace parallel {
inline namespace v1 {

  class exception_list : public exception
  {
    public:
      typedef unspecified iterator;

    size_t size() const noexcept;
    iterator begin() const noexcept;
    iterator end() const noexcept;
```

§ 3.2

```
const char* what() const noexcept override;
};
}
}
}
```

¹ The class exception_list owns a sequence of exception_ptr objects. The parallel algorithms may use the exception_list to communicate uncaught exceptions encountered during parallel execution to the caller of the algorithm.

² The type exception list::iterator shall fulfill the requirements of ForwardIterator.

```
3 size_t size() const noexcept;
```

- ⁴ Returns: The number of exception_ptr objects contained within the exception_list.
- ⁵ *Complexity:* Constant time.
- 6 iterator begin() const noexcept;
 - ⁷ Returns: An iterator referring to the first exception_ptr object contained within the exception_list.
- 8 iterator end() const noexcept;
 - ⁹ *Returns:* An iterator that is past the end of the owned sequence.
- 10 const char* what() const noexcept override;
 - 11 Returns: An implementation-defined NTBS.

§ 3.2

4 Parallel algorithms

[parallel.alg]

4.1 In general

[parallel.alg.general]

This clause describes components that C++ programs may use to perform operations on containers and other sequences in parallel.

4.1.1 Requirements on user-provided function objects

[parallel.alg.general.user]

¹ Function objects passed into parallel algorithms as objects of type BinaryPredicate, Compare, and BinaryOperation shall not directly or indirectly modify objects via their arguments.

4.1.2 Effect of execution policies on algorithm execution [parallel.alg.general.exec]

- ¹ Parallel algorithms have template parameters named ExecutionPolicy which describe the manner in which the execution of these algorithms may be parallelized and the manner in which they apply the element access functions.
- ² The invocations of element access functions in parallel algorithms invoked with an execution policy object of type sequential_execution_policy execute in sequential order in the calling thread.
- ³ The invocations of element access functions in parallel algorithms invoked with an execution policy object of type parallel_execution_policy are permitted to execute in an unordered fashion in unspecified threads, and indeterminately sequenced within each thread. [*Note:* It is the caller's responsibility to ensure correctness, for example that the invocation does not introduce data races or deadlocks. *end note*]

```
[ Example:
```

```
using namespace std::experimental::parallel;
int a[] = {0,1};
std::vector<int> v;
for_each(par, std::begin(a), std::end(a), [&](int i) {
   v.push_back(i*2+1);
});
```

The program above has a data race because of the unsynchronized access to the container v. — *end example*]

[Example:

```
using namespace std::experimental::parallel;
std::atomic<int> x = 0;
int a[] = {1,2};
for_each(par, std::begin(a), std::end(a), [&](int n) {
   x.fetch_add(1, std::memory_order_relaxed);
   // spin wait for another iteration to change the value of x
   while (x.load(std::memory_order_relaxed) == 1) { }
});
```

The above example depends on the order of execution of the iterations, and is therefore undefined (may deadlock). — *end example*]

[Example:

§ 4.1.2

```
using namespace std::experimental::parallel;
int x=0;
std::mutex m;
int a[] = {1,2};
for_each(par, std::begin(a), std::end(a), [&](int) {
    m.lock();
    ++x;
    m.unlock();
});
```

The above example synchronizes access to object x ensuring that it is incremented correctly. — $end\ example\ 1$

⁴ The invocations of element access functions in parallel algorithms invoked with an execution policy of type parallel_vector_execution_policy are permitted to execute in an unordered fashion in unspecified threads, and unsequenced with respect to one another within each thread. [*Note:* This means that multiple function object invocations may be interleaved on a single thread. — end note]

[*Note:* This overrides the usual guarantee from the C++ standard, Section 1.9 [intro.execution] that function executions do not interleave with one another. — *end note*]

Since parallel_vector_execution_policy allows the execution of element access functions to be interleaved on a single thread, synchronization, including the use of mutexes, risks deadlock. Thus the synchronization with parallel vector execution policy is restricted as follows:

A standard library function is *vectorization-unsafe* if it is specified to synchronize with another function invocation, or another function invocation is specified to synchronize with it, and if it is not a memory allocation or deallocation function. Vectorization-unsafe standard library functions may not be invoked by user code called from parallel_vector_execution_policy algorithms.

[*Note:* Implementations must ensure that internal synchronization inside standard library routines does not induce deadlock. — *end note*]

[Example:

```
using namespace std::experimental::parallel;
int x=0;
std::mutex m;
int a[] = {1,2};
for_each(par_vec, std::begin(a), std::end(a), [&](int) {
    m.lock();
    ++x;
    m.unlock();
});
```

The above program is invalid because the applications of the function object are not guaranteed to run on different threads. — *end example*]

[Note: The application of the function object may result in two consecutive calls to m.lock on the same thread, which may deadlock. — $end\ note$]

[Note: The semantics of the parallel_execution_policy or the parallel_vector_execution_policy invocation allow the implementation to fall back to sequential execution if the system cannot parallelize an algorithm invocation due to lack of resources. — $end\ note$]

⁶ Algorithms invoked with an execution policy object of type execution_policy execute internally as if invoked with the contained execution policy object.

§ 4.1.2

7 The semantics of parallel algorithms invoked with an execution policy object of implementation-defined type are implementation-defined.

4.1.3 ExecutionPolicy algorithm overloads

[parallel.alg.overloads]

- ¹ The Parallel Algorithms Library provides overloads for each of the algorithms named in Table 1, corresponding to the algorithms with the same name in the C++ Standard Algorithms Library. For each algorithm in Table 1, if there are overloads for corresponding algorithms with the same name in the C++ Standard Algorithms Library, the overloads shall have an additional template type parameter named ExecutionPolicy, which shall be the first template parameter. In addition, each such overload shall have the new function parameter as the first function parameter of type ExecutionPolicy&&.
- ² Unless otherwise specified, the semantics of ExecutionPolicy algorithm overloads are identical to their overloads without.
- Parallel algorithms shall not participate in overload resolution unless is_execution_policy<decay_t<ExecutionPolicy>>::value is true.

Table 1 — Table of parallel algorithms

<u> </u>	dianci digoritimis	
adjacent_find	all_of	any_of
copy_if	copy_n	count
equal	exclusive_scan	fill
find	find_end	find_first_of
find_if_not	for_each	for_each_n
generate_n	includes	inclusive_scan
inplace_merge	is_heap	is_heap_until
is_sorted	is_sorted_until	${\tt lexicographical_compare}$
merge	min_element	minmax_element
move	none_of	nth_element
partial_sort_copy	partition	partition_copy
remove	remove_copy	remove_copy_if
replace	replace_copy	replace_copy_if
reverse	reverse_copy	rotate
search	search_n	set_difference
$\verb"set_symmetric_difference"$	set_union	sort
stable_sort	swap_ranges	transform
uninitialized_copy_n	$uninitialized_fill$	uninitialized_fill_n
unique_copy		
	adjacent_find copy_if equal find find_if_not generate_n inplace_merge is_sorted merge move partial_sort_copy remove replace reverse search set_symmetric_difference stable_sort uninitialized_copy_n	adjacent_find all_of copy_if copy_n equal exclusive_scan find find_end find_if_not for_each generate_n includes inplace_merge is_heap is_sorted is_sorted_until merge min_element move none_of partial_sort_copy partition remove remove_copy replace replace_copy reverse reverse_copy search search_n set_symmetric_difference set_union stable_sort swap_ranges uninitialized_copy_n uninitialized_fill unique_copy

[Note: Not all algorithms in the Standard Library have counterparts in Table 1. — end note]

4.2 Definitions

[parallel.alg.defns]

```
Define GENERALIZED_SUM(op, a1, ..., aN) as follows:
```

```
a1 when N is 1
```

— op(GENERALIZED_SUM(op, b1, ..., bK), GENERALIZED_SUM(op, bM, ..., bN)) where

— b1, ..., bN may be any permutation of a1, ..., aN and

 $-- 1 < K+1 = M \le N.$

a1 when N is 1

§ 4.2

Define <code>GENERALIZED_NONCOMMUTATIVE_SUM(op, a1, ..., aN)</code> as follows:

```
— op(GENERALIZED\_NONCOMMUTATIVE\_SUM(op, a1, ..., aK), GENERALIZED\_NONCOMMUTATIVE\_SUM(op, aM, ..., aN) where 1 < K+1 = M \leq N.
```

4.3 Non-Numeric Parallel Algorithms

[parallel.alg.ops]

4.3.1 Header <experimental/algorithm> synopsis

[parallel.alg.ops.synopsis]

4.3.2 For each

4

[parallel.alg.foreach]

- Effects: Applies f to the result of dereferencing every iterator in the range [first,last).
 [Note: If the type of first satisfies the requirements of a mutable iterator, f may apply nonconstant functions through the dereferenced iterator. end note]
- ³ Complexity: Applies f exactly last first times.
- ⁴ *Remarks:* If f returns a result, the result is ignored.
- ⁵ *Notes:* Unlike its sequential form, the parallel overload of for_each does not return a copy of its Function parameter, since parallelization may not permit efficient state accumulation.
- ⁶ Requires: Unlike its sequential form, the parallel overload of for_each requires Function to meet the requirements of CopyConstructible.

 $\S~4.3.2~~14$

⁸ Requires: Function shall meet the requirements of MoveConstructible [Note: Function need not meet the requirements of CopyConstructible. — end note]

- ⁹ Effects: Applies f to the result of dereferencing every iterator in the range [first,first + n), starting from first and proceeding to first + n 1. [Note: If the type of first satisfies the requirements of a mutable iterator, f may apply nonconstant functions through the dereferenced iterator. end note]
- 10 Returns: first + n for non-negative values of n and first for negative values.
- 11 *Remarks:* If f returns a result, the result is ignored.

- 13 Effects: Applies f to the result of dereferencing every iterator in the range [first,first + n), starting from first and proceeding to first + n 1. [Note: If the type of first satisfies the requirements of a mutable iterator, f may apply nonconstant functions through the dereferenced iterator. end note]
- 14 Returns: first + n for non-negative values of n and first for negative values.
- 15 *Remarks:* If f returns a result, the result is ignored.
- Notes: Unlike its sequential form, the parallel overload of for_each_n requires Function to meet the requirements of CopyConstructible.

4.4 Numeric Parallel Algorithms

[parallel.alg.numeric]

4.4.1 Header <experimental/numeric> synopsis

[parallel.alg.numeric.synopsis]

§ 4.4.1 15

```
OutputIterator
        exclusive scan(InputIterator first, InputIterator last,
                        OutputIterator result,
                        T init);
    template<class InputIterator, class OutputIterator,
             class T, class BinaryOperation>
      OutputIterator
        exclusive scan(InputIterator first, InputIterator last,
                        OutputIterator result,
                        T init, BinaryOperation binary op);
    template<class InputIterator, class OutputIterator>
      OutputIterator
        inclusive scan(InputIterator first, InputIterator last,
                        OutputIterator result);
    template<class InputIterator, class OutputIterator,
             class BinaryOperation>
      OutputIterator
        inclusive scan(InputIterator first, InputIterator last,
                        OutputIterator result,
                        BinaryOperation binary op);
    template<class InputIterator, class OutputIterator,
             class BinaryOperation, class T>
      OutputIterator
        inclusive scan(InputIterator first, InputIterator last,
                        OutputIterator result,
                        BinaryOperation binary op, T init);
4.4.2 Reduce
                                                                    [parallel.alg.reduce]
        template<class InputIterator>
             typename iterator traits<InputIterator>::value type
               reduce(InputIterator first, InputIterator last);
 <sup>2</sup> Effects: Same as reduce(first, last, typename iterator_traits<InputIterator>::value_type{}).
        template<class InputIterator, class T>
             T reduce(InputIterator first, InputIterator last, T init);

<sup>4</sup> Effects: Same as reduce(first, last, init, plus⇔()).
```

1

§ 4.4.2

```
template<class InputIterator, class T, class BinaryOperation>
T reduce(InputIterator first, InputIterator last, T init,
BinaryOperation binary_op);
```

- 6 Returns: GENERALIZED_SUM(binary_op, init, *first, ..., *(first + (last first) 1)).
- ⁷ Requires: binary_op shall not invalidate iterators or subranges, nor modify elements in the range [first,last).
- 8 Complexity: O(last first) applications of binary op.
- ⁹ *Notes:* The primary difference between reduce and accumulate is that the behavior of reduce may be non-deterministic for non-associative or non-commutative binary op.

4.4.3 Exclusive scan

[parallel.alg.exclusive.scan]

- 4 Effects: Assigns through each iterator i in [result, result + (last first)) the value of GENERALIZED NONCOMMUTATIVE SUM(binary op, init, *first, ..., *(first + (i - result) - 1)).
- ⁵ Returns: The end of the resulting range beginning at result.
- Requires: binary_op shall not invalidate iterators or subranges, nor modify elements in the ranges [first,last) or [result,result + (last first)).
- ⁷ Complexity: O(last first) applications of binary_op.
- ⁸ Notes: The difference between exclusive_scan and inclusive_scan is that exclusive_scan excludes the ith input element from the ith sum. If binary_op is not mathematically associative, the behavior of exclusive_scan may be non-deterministic.

4.4.4 Inclusive scan

[parallel.alg.inclusive.scan]

§ 4.4.4 17

- ⁴ Effects: Assigns through each iterator i in [result, result + (last first)) the value of GENERALIZED_NONCOMMUTATIVE_SUM(binary_op, *first, ..., *(first + (i - result))) or GENERALIZED_NONCOMMUTATIVE_SUM(binary_op, init, *first, ..., *(first + (i - result))) if init is provided.
- ⁵ Returns: The end of the resulting range beginning at result.
- 6 Requires: binary_op shall not invalidate iterators or subranges, nor modify elements in the ranges [first,last) or [result,result + (last - first)).
- ⁷ Complexity: O(last first) applications of binary_op.
- ⁸ Notes: The difference between exclusive_scan and inclusive_scan is that inclusive_scan includes the ith input element in the ith sum. If binary_op is not mathematically associative, the behavior of inclusive_scan may be non-deterministic.

§ 4.4.4 18