experience what's inside"

## WGLISNUEANDEXCUISNUESEAN

SSG DPD , Nikolay Panchenko

## Legal Disclaimer \& Optimization Notice


#### Abstract

INFORMATION IN THIS DOCUMENT IS PROVIDED "AS IS". NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS DOCUMENT. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THIS INFORMATION INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.


Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Copyright © 2016, Intel Corporation. All rights reserved. Intel, Pentium, Xeon, Xeon Phi, Core, VTune, Cilk, and the Intel logo are trademarks of Intel Corporation in the U.S. and other countries.

## Optimization Notice

Intel's compilers may or may not optimize to the same degree for non-Intel microprocessors for optimizations that are not unique to Intel microprocessors. These optimizations include SSE2, SSE3, and SSSE3 instruction sets and other optimizations. Intel does not guarantee the availability, functionality, or effectiveness of any optimization on microprocessors not manufactured by Intel. Microprocessor-dependent optimizations in this product are intended for use with Intel microprocessors. Certain optimizations not specific to Intel microarchitecture are reserved for Intel microprocessors. Please refer to the applicable product User and Reference Guides for more information regarding the specific instruction sets covered by this notice.

Notice revision \#20110804

## Agenda

- Scan
- Motivating examples
- Inclusive scan
- Syntax
- Semantics
- Examples
- Exclusive scan
- Syntax
- Semantics
- Examples
- User-Defined Scan
- Vectorizer features in 18.0
- Explicit syntax for histogram-like pattern
- Explicit syntax for compress and expand -like patterns
- Explicit syntax for conditional lastprivates
- Explicit vectorization of loops with breaks


## Motivating examples

```
Inclusive scan
x = 0;
for (i=0; i < n; ++i) \{
    \(x+=A[i] ;\)
    \(B[i]=x\);
\}
\(B[j]=A[0]+A[1]+\ldots+A[j-1]+A[j] ;\)
```

C++17:
template< class ExecutionPolicy, class ForwardIt1 , class ForwardIt2,
class BinaryOperation, class T >
Forwardlt2
inclusive_scan(ExecutionPolicy\&\& policy, Forwar dlt1 first, Forwardlt1 last, ForwardIt2
d_first, BinaryOperation binary_op, T init );

```
Exclusive scan
x = 0;
for (i=0; i < n; ++i) \{
    \(\mathrm{B}[i]=\mathrm{x}\);
    \(\mathrm{x}+=\mathrm{A}[\mathrm{i}]\);
\}
\(\mathrm{B}[0]=0\);
\(B[j]=A[0]+A[1]+\ldots+A[j-1], j>0 ;\)
```

C++17:
template< class ExecutionPolicy, class ForwardIt1
, class Forwardlt2,
class T, class BinaryOperation >
Forwardlt2
exclusive_scan(ExecutionPolicy\&\& policy,
Forwardlt1 first, Forwardlt1 last, Forwardlt2
d_first, T init, BinaryOperation binary_op );

## Problems with automatic vectorization and/or parallelization

- In general, it's impossible for (i = 0; i < n; ++i) \{
to correctly identify synchronization point
x += C[i];
foo(x, A[i]);
$B[i]=x ;$
\}
- In general, it's impossible to transform exclusive scan to inclusive scan


## Inclusive scan definition

Definition of inclusive scan, of a variable $X$ within the loop L : if two non-empty disjoint statement sequences $\mathbf{S 1}$ and $\mathbf{S} 2$ of $\mathbf{L}$ can be selected, such that

- S1 contains all definitions of $X$; lexically last such statement is W,
- S1 also contains all statements lexically prior to W that directly or indirectly use $X$,
- Any definition in S1 $X$ may only be used in $\mathbf{S 1}$ except the definition of $X$ in the statement $\mathbf{W}$,
- Direct or indirect use of $X$ within $S 1$ must not appear in any condition expression,
- S2 consists of all statements lexically following W.


## Examples for the definition



## Proposed syntax for inclusive scan

new clause: scan(sc : item-list)

- Loop level clause, which is used with \#pragma omp [simd | parallel for | ...]
- sc: scan-combiner is any built-in binary operation (+, -, *, ...) or UDS
- item-list contains one or more scalar variables
new pragma: \#pragma omp inclusive_scan(item-list)
- "partition" pragma, which is used within the loop
- specifies a boundary between definitions and uses
- item-list contains one or more variables that are listed in the scan clause
- item-list cannot contain variables that are not listed in the scan clause
- a variable from item-list must not be used in any other construct
- the pragma must not be used in nested loops
- each list item from scan must be specified in only one inclusive_scan (or exclusive_scan)
- loop may contain multiple instances of inclusive_scan

```
Aligned with reduction/in_reduction(in OpenMP 5.0 Preview 2 (TR6))
    specification.
```


## Use of the syntax for simple case

```
x = 0;
#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
    x += A[i];
#pragma omp inclusive_scan(x)
    B[i] = x;
}
```

```
\(x=0\);
for (i \(=0 ; i<n ; i+=4) \quad\{\)
    \(\mathrm{VX}=[0,0,0,0]\);
    \(\mathrm{VX}=\mathrm{VX}+\mathrm{A}[\mathrm{i}+3: i] ;\)
    rvx \(=[v x 3+v x 2+v x 1+v x 0+x\),
        \(\mathrm{vx} 2+\mathrm{vx} 1+\mathrm{vx} 0+\mathrm{x}\),
        \(\mathrm{Vx} 1+\mathrm{Vx} 0+\mathrm{x}\),
        \(\mathrm{Vx} 0+\mathrm{x}]\);
    \(B[i+3: i]=r v x ;\)
    \(\mathrm{x}=\mathrm{rvx} 3 ;\)
\}
```


## Proposed semantics for inclusive scan

For the scan construct, a private copy $v x$ of each list item is created, one for each SIMD lane as if the private clause had been used. On each iteration of the loop the private copy is initialized with identity value.
Any use of the list item prior to inclusive_scan construct as if it is private. A second private copy $r v x$ for each list item is created at inclusive_scan construct. On each iteration of the loop each SIMD lane of $r v x, r v x_{i}, i \in$ $\{0, \ldots, v l-1\}$, is initialized as

$$
r v x_{i}=\operatorname{sc}\left(\ldots\left(\operatorname{sc}\left(s c\left(\text { original }_{\text {item }}, v x_{0}\right), v x_{1}\right) \ldots v x_{i}\right), \forall i,\right.
$$

Any use of the list item after inclusive_scan construct as if it is second private. At the end of each iteration the original list item is updated by combining the original list item with the all lanes of the private copy using the combiner of the specified scan-combiner.

## Use of the syntax for simple case

For the scan construct, a private copy $v x$ of each list item is created, one for each SIMD lane as if the private clause had been used. On each iteration of the loop the private copy is initialized with identity value.
construct as if it is private.
A second private copy rvx for each list item is created at inclusive_scan construct. On each iteration of the loop each SIMD lane of $r v x, r v x_{i}, i \in\{0, \ldots, v l-1\}$, is
initialized as
$r v x_{i}=\operatorname{sc}\left(\ldots\left(\operatorname{sc}\left(\operatorname{sc}\left(\right.\right.\right.\right.$ original $\left.\left.\left._{i t e m}, v x_{0}\right), v x_{1}\right) \ldots v x_{i}\right), \forall i$,
Any use of the list item after inclusive_scan construct as
if it is second private.
At the end of each iteration the original list item is
updated by $r v x_{v l-1}$ which corresponds to the
lexicographically last executed iteration of the loop.
\}

```
x = 0;
```

x = 0;
\#pragma omp simd scan(+: x)
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
for (i = 0; i < n; ++i) {
x += A[i];
x += A[i];
\#pragma omp inclusive_scan(x)
\#pragma omp inclusive_scan(x)
B[i] = x;
B[i] = x;
}
}
for (i = 0; i < n; i += 4) {
for (i = 0; i < n; i += 4) {
vx = [0, 0, 0, 0];
vx = [0, 0, 0, 0];
\}

```
\}
```

\}

```

\section*{Use of the syntax for simple case}

For the scan construct, a private copy \(v x\) of each list item is created, one for each SIMD lane as if the private clause had been used. On each iteration of the loop the private copy is initialized with identity value.
Any use of the list item prior to inclusive_scan construct as if it is private.
inclusive_scan construct. On each iteration of the loop each SIMD lane of \(r v x, r v x_{i}, i \in\{0, \ldots, v l-1\}\), is
initialized as
\(r v x_{i}=s c\left(\ldots\left(s c\left(s c\left(\right.\right.\right.\right.\) original \(\left.\left.\left._{\text {item }}, v x_{0}\right), v x_{1}\right) \ldots v x_{i}\right), \forall i\),
Any use of the list item after inclusive_scan construct
as if it is second private.
At the end of each iteration the original list item is
updated by \(r v x_{v l-1}\) which corresponds to the
lexicographically last executed iteration of the loop.
```

x = 0;
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
x += A[i];
\#pragma omp inclusive_scan(x)
B[i] = x;
}
for (i = 0; i < n; i += 4) {
vx = vx + A[i + 3:i];

```
\}

\section*{Use of the syntax for simple case}

For the scan construct, a private copy \(v x\) of each list item is created, one for each SIMD lane as if the private clause had been used. On each iteration of the loop the private copy is initialized with identity value.
Any use of the list item prior to inclusive_scan
construct as if it is private.
A second private copy \(r v x\) for each list item is created at inclusive_scan construct. On each iteration of the loop
each SIMD lane of \(r v x, r v x_{i}, i \in\{0, \ldots, v l-1\}\), is
initialized as
\(r v x_{i}=\operatorname{sc}\left(\ldots\left(\operatorname{sc}\left(s c\left(\right.\right.\right.\right.\) original \(\left.\left.\left._{\text {item }}, v x_{0}\right), v x_{1}\right) \ldots v x_{i}\right), \forall i\), Any use of the list item
as if it is second private
At the end of each iteration the original list item is
updated by \(r v x_{v l-1}\) which corresponds to the
lexicographically last executed iteration of the loop.
```

```
x = 0;
```

```
x = 0;
#pragma omp simd scan(+: x)
#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
for (i = 0; i < n; ++i) {
    x += A[i];
    x += A[i];
#pragma omp inclusive_scan(x)
#pragma omp inclusive_scan(x)
    B[i] = x;
    B[i] = x;
}
}
for (i = 0; i < n; i += 4) {
for (i = 0; i < n; i += 4) {
    rvx = [vx3 + vx2 + vx1 + vx0 + x,
    rvx = [vx3 + vx2 + vx1 + vx0 + x,
            vx2 + vx1 + vx0 + x,
            vx2 + vx1 + vx0 + x,
            vx1 + vx0 + x,
            vx1 + vx0 + x,
                        vx0 + x];
                        vx0 + x];
}
```

```
}
```

```

\section*{Use of the syntax for simple case}

For the scan construct, a private copy \(v x\) of each list item is created, one for each SIMD lane as if the private clause had been used. On each iteration of the loop the private copy is initialized with identity value.
Any use of the list item prior to inclusive_scan
construct as if it is private.
A second private copy rvx for each list item is created at inclusive_scan construct. On each iteration of the loop each SIMD lane of \(r v x, r v x_{i}, i \in\{0, \ldots, v l-1\}\), is
initialized as

Any use of the list item after inclusive_scan construct as if it is second private.
At the end of each iteration the original list item is
updated by \(r v x_{v l-1}\) which corresponds to the
lexicographically last executed iteration of the loop.
```

x = 0;
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
x += A[i];
\#pragma omp inclusive_scan(x)
B[i] = x;
}
for (i = 0; i < n; i += 4) {
B[i + 3:i] = rvx;

```
\}

\section*{Use of the syntax for simple case}

For the scan construct, a private copy \(v x\) of each list item is created, one for each SIMD lane as if the private clause had been used. On each iteration of the loop the private copy is initialized with identity value.
Any use of the list item prior to inclusive_scan
construct as if it is private.
A second private copy rox for each list item is created ai inclusive_scan construct. On each iteration of the loop each SIMD lane of \(r v x, r v x_{i}, i \in\{0, \ldots, v l-1\}\), is
initialized as
Any use of the list item after inclusive_scan construct
At the end of each iteration the original list item is updated by \(r v x_{v l-1}\) which corresponds to the lexicographically last executed iteration of the loop.
```

x = 0;
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
x += A[i];
\#pragma omp inclusive_scan(x)
B[i] = x;
}
for (i = 0; i < n; i += 4) {
x = rvx3;
}

```

\section*{Programmer can cheat}
```

x = 1;
\#pragma omp simd scan(*: x)
for (i = 0; i < n; ++i) {
x += i;
x += A[i];
\#pragma omp inclusive_scan(x)
B[i] = x;
}

```
- This is legal
- Execution will be different to scalar execution
- Aligned with reduction/in_reduction (in TR6) specification

\section*{Use of the syntax with conditional updates}
```

x = 0;
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
if (i \& 5) {
x += A[i];
}
else {
x += 2 * A[i];
}
\#pragma omp inclusive_scan(x)
B[i] = x;
}

```

\section*{inclusive_scan under condition}
```

x = 0;
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
if (i \& 5) {
x += A[i];
\#pragma omp inclusive_scan(x)
B[i] = x;
}
}

```
```

x = 0;
for (i = 0; i < n; i += 4) {
VX = [0, 0, 0, 0];
m = ([i + 3, i + 2, i + 1, i] \& [5, 5, 5, 5])
!=0;
VX = vx + @m A[i + 3:i]@m;
// for all lanes of the rvx
rvx = [vx3 + vx2 + vx1 + vx0 + x,
vx2 + vx1 + vx0 + x,
vx1 + vx0 + x,
vx0 + x];
B[i + 3:i] = @m rvx;
x = rvx3;
}

```

\section*{Use of the syntax for two scans}
```

x = 0;
y = 0;
\#pragma omp simd scan(+: x, y)
for (i = 0; i < n; ++i) {
x += A[i];
\#pragma omp inclusive_scan(x)
B[i] = x;
y += B[i];
\#pragma omp inclusive_scan(y)
C[i] = y;
}

```
```

x = 0;
y = 0;
for (i = 0; i < n; i += 4) {
vx = [0, 0, 0, 0];
VY = [0, 0, 0, 0];
VX = vx + A[i + 3:i];
rvx}=[vx3+vx2 + vx1 + vx0 + x
vx2 + vx1 + vx0 + x,
vx1 + vx0 + x,
vx0 + x];
B[i + 3:i] = rvx;
x = rxv3;
VY = vY + B[i + 3:i];
rvy = [vy3 + vy2 + vy1 + vy0 + y,
vy2 + vy1 + vy0 + y,
vy1 + vy0 + y,
vy0 + y];
C[i + 3:i] = rvy;
y = rvy3;
}

```

\section*{Non-conformant scan with outer loop}
```

x = 0;
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
x += A[i];
for (j = 0; j < m; ++j) {
x += C[j];
\#pragma omp inclusive_scan(x)
B[j] = x;
}
}

```
In reality this loop is:
\#pragma omp simd \(\operatorname{scan}(+: x)\)
for ( \(i=0\); \(i<n\); ++i) \{
    x \(+=\mathrm{A}[\mathrm{i}] ;\)
    // j loop
    \(\mathrm{x}+=\mathrm{C}[0]\);
\#pragma omp inclusive_scan(x)
    \(\mathrm{B}[0]=\mathrm{x}\);
    \(\mathrm{x}+=\mathrm{C}[1]\);
\#pragma omp inclusive_scan(x)
    \(B[1]=x\);
    \(\mathrm{x}+=\mathrm{C}[\mathrm{m}-1]\);
\#pragma omp inclusive_scan(x)
    \(B[m-1]=x ;\)
\}
which fall into non - conformant case.
So \#pragma omp inclusive_scan must be used in the same loop level as corresponding scan clause.

\section*{Conformant scan with outer loop}
```

x = 0;
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
x += A[i];
for (j = 0; j < m; ++j) {
x += C[j];
}
\#pragma omp inclusive_scan(x)
B[j] = x;
}

```
```

x = 0;
for (i = 0; i < n; i += 4) {
VX = [0, 0, 0, 0];
VX = VX + A[i];
for (j = 0; j < m; ++j) {
vx += broadcast(C[j]);
}
rvx}=[vx3+vx2+vx1 + vx0 + x,
vx2 + vx1 + vx0 + x,
vx1 + vx0 + x,
vx0 + x];
B[i + 3:i] = rvx;
x = rvx3;
}

```

\section*{Exclusive scan}
```

Simple exclusive scan
x = 0;
for (i = 0; i < n; ++i) {
B[i] = x;
x += A[i];
}
B[0] = 0;
B[j] = A[0]+A[1]+...+A[j-1], j > 0;

```

Simple exclusive scan
for (i = 0; i < n; ++i) {
    x += A[i]; // With special
processing
    B[i] = x;
    _%+- त[i];
}
```

Required transformation (exclusive -> inclusive scan)

```
```

x = 0;

```
```

```
x = 0;
```


## Exclusive scan definition

Definition of exclusive scan, of a variable $X$ within the loop L : if two disjoint non-empty statement sequences $\mathbf{S 1}$ and $\mathbf{S} 2$ of $\mathbf{L}$ can be selected, such that

- Lexically last statement of $\mathbf{S 1}$ precedes to lexically first statement of $\mathbf{S 2}$
- S2 must contain at least one write to X
- $\mathbf{S 1}$ must not have any writes to $X$
- Any statement in $L$ must belong to $\mathbf{S 1}$ or $\mathbf{S} 2$
- S1 and S2 are independent, except dependencies to/from $X$
- Direct or indirect use of $X$ within $\mathbf{S} 2$ must not appear in any condition expression


## Examples for the definition



## Proposed syntax for exclusive scan

scan(sc: item-list)

- Loop level clause, which is used with \#pragma omp [simd]
- sc: scan-combiner is any built-in binary operation (+, -, *, ...) or UDS
- item-list contains one or more scalar variables
new directive: \#pragma omp exclusive_scan(item-list)
- "partition" pragma, which is used within the loop and allows to move block of statement before the pragma at the end of the loop or move block of statements after the pragma at the beginning of the loop
- item-list contains one or more variables that are listed in the scan clause
- item-list cannot contain variables that are not listed in the scan clause
- a variable from item-list must not be used in any other construct
- pragma must not be used under conditions or nested loops
- loop must have at most one instance of exclusive_scan
exclusive_scan pragma specifies end point of the block of statements that can be moved at the end of the loop.
inclusive_scan pragma specifies place where running-reduction vector can be safely computed


## Use of the syntax for simple case

```
x = 0;
#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
    B[i] = x;
#pragma omp exclusive_scan(x)
    x += A[i];
}
```



```
x = 0;
```

x = 0;
\#pragma omp simd scan(+: x)
\#pragma omp simd scan(+: x)
for (i = 0; i < n; ++i) {
for (i = 0; i < n; ++i) {
x += A[i];
x += A[i];
\#pragma omp
\#pragma omp
__inclusive_scan_shifted_by_one(x)
__inclusive_scan_shifted_by_one(x)
B[i] = x;
B[i] = x;
}

```
}
```

inclusive_scan(x) =
$[x 3+x \overline{2}+x 1+x 0+x$ _init,
$x 2+x 1+x 0+x \_i n i t$,
$\mathrm{x} 1+\mathrm{x} 0$ + x _init,
$x 0+x$ _-init];

## Proposed semantics for exclusive scan

For the scan construct, a private copy $v x$ of each list item is created, one for each SIMD lane as if the private clause has been used. On each iteration of the loop the private copy is initialized with identity value.
Loop execution is changed as if each list item is transformed to inclusive scan.
A second private copy $r v x$ for each list item is created at exclusive_scan construct. On each iteration of the loop each SIMD lane of $r v x, r v x_{i}, i \in$ $\{0, \ldots, v l-1\}$, is initialized as

$$
\begin{gathered}
r v x_{0}=\text { original }_{\text {item }} \\
r v x_{i}=\operatorname{sc}\left(\ldots\left(\operatorname{sc}\left(\operatorname{sc}\left(\text { original }_{\text {item }}, v x_{0}\right), v x_{1}\right) \ldots v x_{i-1}\right), \forall i \in\{1, \ldots, v l-1\},\right.
\end{gathered}
$$

Any use of the list item prior to exclusive_scan construct as if it is second private.
At the end of each iteration the original list item is updated by combining the original list item with the all lanes of the private copy using the combiner of the specified scan-combiner.

## Find the right place for exclusive_scan

```
x = 0;
#pragma omp simd scan(+:x)
for (i = 0; i < n; ++i) {
    if (A[i] > 0) {
        C[i] = x;
    }
#pragma omp exclusive_scan(x)
    D[i] = C[i];
#pragma omp exclusive_scan(x)
    if (A[i] < 0) {
        x += A[i];
    }
}
```


## Inclusive and Exclusive scans in the same loop

```
x = 0;
y = 0;
#pragma omp simd scan(+:x, y)
for (i = 0; i < n; ++i) {
    B[i] = x;
#pragma omp exclusive_scan(x)
    x += A[i];
    y += A[i];
#pragma omp inclusive_scan(y)
    C[i] = y;
}
```


## User-Defined Scan (UDS)

new pragma: \#pragma omp declare scan(scan-identifier : typename-list : combiner) [initializerclause] new-line

- Same as declare reduction except for scan keyword,
- scan-identifier is either a base language identifier or one of the following operations: +, -, *, \&, |, ^, \&\& and ||
- typename-list is a list of type names
- combiner is an expression
- initializer-clause is initializer(initializer-expr) where initializer-expr is omp_priv = initializer or function-name(argument-list).


## Explicit syntax for histogram-like pattern

```
for (i = 0; i < n; ++i) {
    ++a[b[i]];
}
```


## Explicit syntax for compress-like pattern

```
for (i = 0; i < n; ++i) {
    if (condition) {
        a[++j] = b[i];
    }
}
```

```
#pragma omp simd
for (i = 0; i < n; ++i) {
    #pragma omp ordered simd monotonic(j:1)
    {
        if (condition) {
        a[++j] = b[i];
        }
    }
```

\}

## Explicit syntax for expand-like pattern

```
for (i = 0; i < n; ++i) {
    if (condition) {
        b[i] = a[j++];
    }
}
```

\#pragma omp simd
for ( $\mathrm{i}=0$; $\mathrm{i}<\mathrm{n}$; ++i) \{
\#pragma omp ordered simd monotonic(j:1)
\{
if (condition) \{
$\mathrm{b}[\mathrm{i}]=\mathrm{a}[\mathrm{j}++]$;
\}
\}

## Explicit syntax for conditional lastprivate

```
for (i = 0; i < n; ++i) {
    if (condition) {
        j = i;
    }
}
```

\#pragma omp simd
lastprivate(conditional: j)
for (i = 0; i < n; ++i) \{
if (condition) \{
j = i;
\}

## Explicit syntax for loops with exits

```
for (i = 0; i < n; ++i) {
    if (condition) {
        j = i;
        break;
    }
}
```

\#pragma omp simd early_exit
lastprivate(conditional: j)
for (i = 0; i < n; ++i) \{
if (condition) \{
j = i;
break;
\}

