Intel® High Level Synthesis Compiler Pro Edition

Reference Manual

Updated for Intel® Quartus® Prime Design Suite: 21.2
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.4.</td>
<td>Loop Unrolling (unrollPragma)</td>
<td>65</td>
</tr>
<tr>
<td>6.5.</td>
<td>Loop Concurrency (max_concurrencyPragma)</td>
<td>66</td>
</tr>
<tr>
<td>6.6.</td>
<td>Loop Iteration Speculation (speculated_iterationsPragma)</td>
<td>67</td>
</tr>
<tr>
<td>6.7.</td>
<td>Loop Pipelining Control (disable_loop_pipeliningPragma)</td>
<td>68</td>
</tr>
<tr>
<td>6.8.</td>
<td>Loop Interleaving Control (max_interleavingPragma)</td>
<td>68</td>
</tr>
<tr>
<td>6.9.</td>
<td>Loop Fusion</td>
<td>71</td>
</tr>
<tr>
<td>6.9.1.</td>
<td>Loop Fusion Control (loop_fusePragma)</td>
<td>73</td>
</tr>
<tr>
<td>6.9.2.</td>
<td>Loop Fusion Exemption (nofusionPragma)</td>
<td>75</td>
</tr>
<tr>
<td>7.</td>
<td>Component Concurrency</td>
<td>76</td>
</tr>
<tr>
<td>7.1.</td>
<td>Serial Equivalence within a Memory Space or I/O</td>
<td>76</td>
</tr>
<tr>
<td>7.2.</td>
<td>Concurrency Control (hls_max_concurrencyAttribute)</td>
<td>76</td>
</tr>
<tr>
<td>7.3.</td>
<td>Component Pipelining Control (hls_disable_component_pipeliningAttribute)</td>
<td>78</td>
</tr>
<tr>
<td>8.</td>
<td>Arbitrary Precision Math Support</td>
<td>79</td>
</tr>
<tr>
<td>8.1.</td>
<td>Declaring ac_intData Types</td>
<td>81</td>
</tr>
<tr>
<td>8.1.1.</td>
<td>Important Usage Information on the ac_intData Type</td>
<td>82</td>
</tr>
<tr>
<td>8.1.2.</td>
<td>Integer Promotion and ac_intData Types</td>
<td>82</td>
</tr>
<tr>
<td>8.1.3.</td>
<td>Debugging Your Use of the ac_intData Type</td>
<td>86</td>
</tr>
<tr>
<td>8.2.</td>
<td>Declaring ac_fixedData Types</td>
<td>86</td>
</tr>
<tr>
<td>8.2.1.</td>
<td>Arbitrary Precision Fixed-Point Literals in Operations</td>
<td>87</td>
</tr>
<tr>
<td>8.3.</td>
<td>Declaring ac_complexData Types</td>
<td>89</td>
</tr>
<tr>
<td>8.4.</td>
<td>AC Data Types and Native Compilers</td>
<td>90</td>
</tr>
<tr>
<td>8.5.</td>
<td>Declaring hls_floatData Types</td>
<td>90</td>
</tr>
<tr>
<td>8.5.1.</td>
<td>Operators and Return Types Supported by the hls_floatData Type</td>
<td>92</td>
</tr>
<tr>
<td>8.5.2.</td>
<td>Additional Data Types Provided By hls_float.h</td>
<td>97</td>
</tr>
<tr>
<td>9.</td>
<td>Component Target Frequency</td>
<td>98</td>
</tr>
<tr>
<td>9.1.</td>
<td>Effects of Specifying Target II and Target f_{MAX}</td>
<td>99</td>
</tr>
<tr>
<td>10.</td>
<td>Systems of Tasks</td>
<td>100</td>
</tr>
<tr>
<td>10.1.</td>
<td>Task Functions</td>
<td>101</td>
</tr>
<tr>
<td>10.2.</td>
<td>Internal Streams</td>
<td>109</td>
</tr>
<tr>
<td>10.3.</td>
<td>System of Tasks Simulation</td>
<td>110</td>
</tr>
<tr>
<td>11.</td>
<td>Libraries</td>
<td>111</td>
</tr>
<tr>
<td>11.1.</td>
<td>Static-Object Libraries</td>
<td>112</td>
</tr>
<tr>
<td>11.2.</td>
<td>Creating a Static-Object Library</td>
<td>113</td>
</tr>
<tr>
<td>11.3.</td>
<td>Creating Objects From HLS Code</td>
<td>114</td>
</tr>
<tr>
<td>11.3.1.</td>
<td>Creating an Object File From HLS Code</td>
<td>114</td>
</tr>
<tr>
<td>11.3.2.</td>
<td>Supported OpenCL Language Constructs</td>
<td>116</td>
</tr>
<tr>
<td>11.4.</td>
<td>Creating Objects From RTL Code</td>
<td>117</td>
</tr>
<tr>
<td>11.4.1.</td>
<td>RTL Modules and the HLS Pipeline</td>
<td>118</td>
</tr>
<tr>
<td>11.4.2.</td>
<td>Creating a Static-Object File from an RTL Module</td>
<td>136</td>
</tr>
<tr>
<td>11.5.</td>
<td>Packaging Object Files Into a Library</td>
<td>137</td>
</tr>
<tr>
<td>12.</td>
<td>Advanced Hardware Synthesis Controls</td>
<td>139</td>
</tr>
<tr>
<td>12.1.</td>
<td>Hardware Implementation Control for Math Functions</td>
<td>139</td>
</tr>
<tr>
<td>12.1.1.</td>
<td>Math Function Hardware Implementation Summary</td>
<td>140</td>
</tr>
<tr>
<td>12.1.2.</td>
<td>The --dsp-mode Command Option</td>
<td>141</td>
</tr>
<tr>
<td>12.1.3.</td>
<td>The ihc::math dsp control Function</td>
<td>141</td>
</tr>
</tbody>
</table>

The Intel® HLS Compiler Pro Edition Reference Manual provides reference information about the features supported by the Intel HLS Compiler Pro Edition. The Intel HLS Compiler is sometimes referred to as the i++ compiler, reflecting the name of the compiler command.

In this publication, `<quartus_installdir>` refers to the location where you installed Intel Quartus® Prime Design Suite.

The default Intel Quartus Prime Design Suite installation location depends on your operating system:

**Windows**

C:\intelFPGA_pro\21.2

**Linux**

/home/<username>/intelFPGA_pro/21.2

About the Intel HLS Compiler Pro Edition Documentation Library

Documentation for the Intel HLS Compiler Pro Edition is split across a few publications. Use the following table to find the publication that contains the Intel HLS Compiler Pro Edition information that you are looking for:

<table>
<thead>
<tr>
<th>Title and Description</th>
<th>PRO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Release Notes</td>
<td>Link</td>
</tr>
<tr>
<td>Get late-breaking information about the Intel HLS Compiler.</td>
<td></td>
</tr>
<tr>
<td>Getting Started Guide</td>
<td>Link</td>
</tr>
<tr>
<td>Get up and running with the Intel HLS Compiler by learning how to initialize your compiler environment and reviewing the various design examples and tutorials provided with the Intel HLS Compiler.</td>
<td></td>
</tr>
<tr>
<td>User Guide</td>
<td>Link</td>
</tr>
<tr>
<td>Provides instructions on synthesizing, verifying, and simulating intellectual property (IP) that you design for Intel FPGA products. Go through the entire development flow of your component from creating your component and testbench up to integrating your component IP into a larger system with the Intel Quartus Prime software.</td>
<td></td>
</tr>
<tr>
<td>Reference Manual</td>
<td>Link</td>
</tr>
<tr>
<td>Provides reference information about the features supported by the Intel HLS Compiler. Find details on Intel HLS Compiler command options, header files, pragmas, attributes, macros, declarations, arguments, and template libraries.</td>
<td></td>
</tr>
<tr>
<td>Best Practices Guide</td>
<td>Link</td>
</tr>
<tr>
<td>Provides techniques and practices that you can apply to improve the FPGA area utilization and performance of your HLS component. Typically, you apply these best practices after you verify the functional correctness of your component.</td>
<td></td>
</tr>
<tr>
<td>Quick Reference</td>
<td>Link</td>
</tr>
<tr>
<td>Provides a brief summary of Intel HLS Compiler declarations and attributes on a single two-sided page.</td>
<td></td>
</tr>
</tbody>
</table>
2. Compiler

2.1. Intel HLS Compiler Pro Edition Command Options

Use the Intel HLS Compiler Pro Edition command options to customize how the compiler performs general functions, customize file linking, or customize compilation.

Table 2. General Command Options

These i++ command options perform general compiler functions.

<table>
<thead>
<tr>
<th>Command Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-h or --help</td>
<td>Instructs the compiler to list all the command options and their descriptions on screen.</td>
</tr>
<tr>
<td>-o &lt;result&gt;</td>
<td>Instructs the compiler to place its output into the &lt;result&gt; executable and the &lt;result&gt;.prj directory. If you do not specify the -o &lt;result&gt; option, the compiler outputs an a.out file for Linux and an a.exe file for Windows. Use the -o &lt;result&gt; command option to specify the name of the compiler output. Example command: i++ -o hlsoutput multiplier.c Invoking this example command creates an hlsoutput executable for Linux and an hlsoutput.exe for Windows in your working directory.</td>
</tr>
<tr>
<td>--debug-log</td>
<td>Instructs the compiler to generate a log file that contains diagnostic information. The debug.log file is in the &lt;result&gt;.prj subdirectory within your current working directory. If your compilation fails, the debug.log file is generated whether you set this option or not.</td>
</tr>
<tr>
<td>-v</td>
<td>Verbose mode that instructs the compiler to display messages describing the progress of the compilation. Example command: i++ -v multiplier.cpp, where multiplier.cpp is the input file.</td>
</tr>
<tr>
<td>--version</td>
<td>Instructs the compiler to display its version information on screen. Command: i++ --version</td>
</tr>
</tbody>
</table>
### Table 3. Command Options that Customize Compilation
These `i++` command options perform compiler functions that impact the translation from source file to object file.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>-c</code></td>
<td>Instructs the compiler to preprocess, parse, and generate object files (.o/.obj) in the current working directory. The linking stage is omitted. Example command: <code>i++ -march=&quot;Arria 10&quot; -c multiplier.c</code>. Invoking this example command creates a <code>multiplier.o</code> file and sets the name of the <code>&lt;result&gt;.prj</code> directory to <code>multiplier.prj</code>. When you later link the <code>.o</code> file, the <code>-o</code> option affects only the name of the executable file. The name of the <code>&lt;result&gt;.prj</code> directory remains unchanged from when the directory name was set by <code>i++ -c</code> command invocation.</td>
</tr>
<tr>
<td><code>--component &lt;components&gt;</code></td>
<td>Allows you to specify a comma-separated list of function names that you want the compiler to synthesize to RTL. Example command: <code>i++ counter.cpp --component count to use this option, your component must be configured with C-linkage using the extern &quot;C&quot; specification. For example: extern &quot;C&quot; int myComponent(int a, int b)</code></td>
</tr>
<tr>
<td><code>-D &lt;macro&gt; [= &lt;val&gt;]</code></td>
<td>Allows you to pass a macro definition (&lt;macro&gt;) and its value (&lt;val&gt;) to the compiler. If you do not specify a value for <code>&lt;val&gt;</code>, its default value will be 1.</td>
</tr>
<tr>
<td><code>-g</code></td>
<td>Generate debug information (default).</td>
</tr>
<tr>
<td><code>-g0</code></td>
<td>Do not generate debug information.</td>
</tr>
<tr>
<td><code>--gcc-toolchain=&lt;GCC_dir&gt;</code></td>
<td>Specifies the path to a GCC installation that you want to use for compilation. This path should be the absolute path to the directory that contains the GCC <code>lib</code>, <code>bin</code>, and include folders. You should not need to use this if you configured your system as described in the Getting Started Guide.</td>
</tr>
<tr>
<td><code>--hyper-optimized-handshaking=[auto/off]</code></td>
<td>This option applies to Intel Agilex™ and Intel Stratix® 10 devices only. Use this option to modify the handshaking protocol used in certain areas of your design. By default, the <code>--hyper-optimized-handshaking</code> option is set to <code>auto</code>. When you enable this optimization, the compiler adds pipeline registers to the handshaking paths of the stallable nodes. This optimization results in a higher $f_{\text{MAX}}$ at the cost of increased area and latency due to the added registers. Disabling this optimization typically decreases area and latency at the cost of lower $f_{\text{MAX}}$. Restriction: This option applies only to designs targeting Intel Agilex and Intel Stratix 10 devices. If you use this option when you target devices other than Intel Agilex and Intel Stratix 10 devices, the compiler exits with an error.</td>
</tr>
<tr>
<td><code>-I &lt;dir&gt;</code></td>
<td>Adds a directory (<code>&lt;dir&gt;</code>) to the end of the include path list.</td>
</tr>
</tbody>
</table>

*continued...*
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
</table>
| -march=[x86-64 | <FPGA_family> | <FPGA_part_number>] | Instructs the compiler to compile the component to the specified architecture or FPGA family. The -march compiler option can take one of the following values:  
  x86-64 | Instructs the compiler to compile the code for an emulator flow.  
  "<FPGA_family>" | Instructs the compiler to compile the code for a target FPGA device family. The <FPGA_family> value can be any of the following device families:  
  • Agilex  
  • Arria10 or "Arria 10"  
  • CycloneV(1)  
  • Cyclone10GX or "Cyclone 10 GX"  
  • Stratix10 or "Stratix 10"  
  Quotation marks are required only if you specify a FPGA family name specifier that contains spaces  
  <FPGA_part_number> | Instructs the compiler to compile the code for a target device. The compiler determines the FPGA device family from the FPGA part number that you specify here.  
  If you do not specify this option, -march=x86-64 is assumed.  
  If the parameter value that you specify contains spaces, surround the parameter value in quotation marks. |
| --quartus-compile                           | Compiles your design with the Intel Quartus Prime compiler. Example command: i++ --quartus-compile <input_files> -march="Arria 10"  
  When you specify this option, the Intel Quartus Prime compiler is run after the RTL is generated. The compiled Intel Quartus Prime project is put in the <result>.prj/quartus directory and a summary of the FPGA resource consumption and maximum clock frequency is added to the high level design reports in the <result>.prj/reports directory.  
  This compilation is intended to estimate the best achievable fMAX for your component. Your component is not expected to cleanly close timing in the reports. |
| --quartus-seed <seed>                       | Specifies the seed numeric value that is used by Intel Quartus Prime project located in the <result>.prj/quartus directory. This seed value is used by the Intel Quartus Prime Fitter for initial placement configuration when optimizing design placement to meet timing requirements (fMAX).  
  For more information about the Fitter seed, refer to Fitter Initial Placement Seed in the "Fitter Settings Reference" section of Intel Quartus Prime Pro Edition User Guide: Design Compilation. |
| --simulator <simulator_name>                | Specifies the simulator you are using to perform verification. This command option can take the following values for <simulator_name>:  
  • modelsim  
  • none  
  If you do not specify this option, --simulator modelsim is assumed.  
  Important: The --simulator command option only works in conjunction with the -march command option. |

(1) Cyclone® V device support is subject to restrictions. For details, refer to Cyclone V Restrictions on page 213.
<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>--simulator none</td>
<td>The --simulator none option instructs the HLS compiler to skip the verification flow and generate RTL for the components without generating the corresponding test bench. If you use this option, the high-level design report (report.html) is generated more quickly but you cannot simulate your design. Without data from simulation, the report must omit verification statistics such as component latency. Example command: i++ -march=&quot;&lt;FPGA_family_or_part_number&gt;&quot; --simulator none multiplier.c</td>
</tr>
<tr>
<td>-ffp-contract=fast</td>
<td>Remove intermediate rounding and conversion when possible, except for code blocks fenced by #pragma clang fp contract(off). To learn more, review the following tutorial: &lt;quartus_installdir&gt;/hls/examples/tutorials/best_practices/floating_point_ops</td>
</tr>
<tr>
<td>-ffp-reassoc</td>
<td>Deprecated: This command option is deprecated and will be removed in a later release of the Intel HLS Compiler. Use the -ffp-reassociate command option instead. Relax the order of floating point arithmetic operations, except for code blocks fenced by #pragma clang fp reassoc(off). To learn more, review the following tutorial: &lt;quartus_installdir&gt;/hls/examples/tutorials/best_practices/floating_point_ops</td>
</tr>
<tr>
<td>-ffp-reassociate</td>
<td>Relax the order of floating point arithmetic operations, except for code blocks fenced by #pragma clang fp reassociate(off) To learn more, review the following tutorial: &lt;quartus_installdir&gt;/hls/examples/tutorials/best_practices/floating_point_ops</td>
</tr>
<tr>
<td>-daz</td>
<td>For double data types only, disable subnormal support in native IEEE-754 double-precision floating-point computations. To learn more, review the following tutorial: &lt;quartus_installdir&gt;/hls/examples/tutorials/best_practices/subnormal_and_rounding</td>
</tr>
</tbody>
</table>
| --rounding=[ieee | faithful] | For double data types only, control rounding scheme for native IEEE-754 double-precision adders, multipliers, and dividers. If you do not specify this option, adders and multipliers use IEEE-754 round to nearest, ties to even (RNE) rounding (0.5 ULP) and dividers use faithful rounding (1 ULP). The --rounding option can take one of the following values:  
  ieee  All adders, multipliers, and dividers use IEEE-754 RNE rounding. IEEE-754 RNE rounding rounds results to the nearest value. If the number falls midway, it is rounded to the nearest value with an even least-significant digit. This is the default rounding mode defined by IEEE 754-2008 standard. IEEE-754 RNE rounding has an accuracy of 0.5 ULP.  
  faithful  All adders, multipliers, and dividers use faithful rounding. Faithful rounding rounds results to either the upper or lower nearest single-precision numbers. Therefore, faithful rounding produces one of two possible values. The choice between the two is not defined. Faithful rounding has a maximum error of one ULP. Errors are not guaranteed to be evenly distributed. Faithful rounding mode is not defined by the IEEE-754 standard. To learn more, review the following tutorial: <quartus_installdir>/hls/examples/tutorials/best_practices/subnormal_and_rounding |
| --clock <clock target>     | Optimizes the RTL for the specified clock frequency or period. The <clock target> value must include a unit.                                                                                                                                                                                                                                                                                                                                                           |

**continued...**
Table 4. **Command Options that Customize File Linking**
These HLS command options specify compiler actions that impact the translation of the object file to the binary or RTL component.

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>option</td>
<td>Description</td>
</tr>
<tr>
<td>-ghdl[=&lt;depth&gt;]</td>
<td>Logs signals when running the verification executable to help you debug the generated RTL. After running the executable, the simulator logs waveforms to the a.prj/verification/vsim.wlf file. Use the optional &lt;depth&gt; attribute to specify how many levels of hierarchy are logged. If you do not specify a value for the &lt;depth&gt; attribute, all signals are logged. Use -ghdl=1 to log only the top-level signals. For details about the ModelSim* waveform, see Debugging during Verification in Intel High Level Synthesis Compiler Pro Edition User Guide.</td>
</tr>
<tr>
<td>-L&lt;dir&gt;</td>
<td>(Linux only) Adds a directory (&lt;dir&gt;) to the end of the search path for the library files.</td>
</tr>
<tr>
<td>-l&lt;library&gt;</td>
<td>(Linux only) Specifies the library file (.a) name when linking the object file to the binary. On Windows, you can list library files (.lib) on the command line without specifying any command options or flags.</td>
</tr>
<tr>
<td>--x86-only</td>
<td>Creates only the testbench executable. The compiler outputs a &lt;result&gt; file for Linux or a &lt;result&gt;.exe file for Windows. The &lt;result&gt;.prj directory and its contents are not created.</td>
</tr>
<tr>
<td>--fpga-only</td>
<td>Creates only the &lt;result&gt;.prj directory and its contents. The testbench executable file (&lt;result&gt;/&lt;result&gt;.exe) is not created. Before you can simulate your hardware from a compilation output that uses this option, you must compile your testbench with the --x86-only option (or as part of a full compilation).</td>
</tr>
</tbody>
</table>
2.2. Using Libraries in Your Component

Use libraries to reuse functions created by you or others without needing to know the function implementation details.

To use the functions in a library, you must have the C/C++ header files (.h or .hpp) for the library available. For object libraries, you must also have the object library archive file (.a on Linux systems or .lib on Windows systems) available.

Any object libraries that you use in your component must be built and used by the same version number Intel FPGA high-level design tool. For example, to compile your component with the Intel HLS Compiler Version 21.2, the libraries included in your component must have been created with a version 21.2 Intel FPGA high-level design tool. If you use a library with a different version, you get a version mismatch error when you compile your component.

To include a library in your component:

1. Review the header files corresponding to the library that you want to include in your component.
   The header file shows you the functions available to call in the library and how to call the functions.

2. Include the header files in your component code.
   For example, #include "primitives.h"

3. Compile your component with the Intel HLS Compiler as follows:
   - For source code (that is, header-only) libraries, there is no additional library file name to specify.
     For example, i++ -march=arria10 MyComponent.cpp
   - For object libraries, ensure that you add the object library archive file name to the i++ command.
     For example, i++ -march=arria10 MyComponent.cpp libprim.a (Linux) or i++ -march=arria10 MyComponent.cpp libprim.lib (Windows).

Related Information
- Libraries on page 111
- Advanced Math Source Code Libraries on page 198
- Arbitrary Precision Math Support on page 79

2.3. Compiler Interoperability

You can compile your testbench code with GCC or Microsoft Visual Studio, but generating RTL and simulation support for your component always requires the Intel HLS Compiler.

The following table shows which parts of your design you can compile with each compiler:
<table>
<thead>
<tr>
<th>Component</th>
<th>GCC/MSVC</th>
<th>i++</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testbench</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Component (emulation)</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Component (RTL)</td>
<td>X</td>
<td></td>
</tr>
</tbody>
</table>

**Restriction:** You cannot use GCC and Microsoft Visual Studio to compile code that uses any Intel HLS Compiler provided header files other than HLS/hls.h. For some arbitrary precision data types, the Intel HLS Compiler provides reference headers that can be compiled GCC and Microsoft Visual Studio along with the FPGA-optimized headers. For details, see Arbitrary Precision Math Support on page 79.

To see what versions of GCC and Microsoft Visual Studio the Intel HLS Compiler supports, see "Intel High Level Synthesis Compiler Prerequisites" in Intel High Level Synthesis Compiler Getting Started Guide.

To use a native compiler (GCC or Microsoft Visual Studio) to compile your Intel HLS Compiler code, you must run your native compiler from a terminal session where you initialized the environment for the Intel HLS Compiler. The initialization script chooses the correct native compiler for your system.

**GCC**

To compile your Intel HLS Compiler code with GCC:

1. Initialize your environment with the Intel HLS Compiler initialization script:

   `<quartus_installdir>/hls/init_hls.sh`

2. Add the path to the Intel HLS Compiler header files to the `g++` command include path with the `-l` command option.
   
   The header files are in the `quartus_installdir/hls/include` directory.

3. Add the path to the HLS emulation library to the linker search path with the `-L` command option.
   
   The emulation library is in the `quartus_installdir/hls/host/linux64/lib` directory.

4. Add the `hls_emul` library to the linker command (that is, specify `-lhls_emul` as a command option).

5. Ensure that you specify the `-std=c++17` option of the `g++` command.
6. If you want to generate debug symbols, specify the `--g` option of the `g++` command.

7. If you are using HLS tasks in a system of tasks (`ihc::launch` and `ihc::collect`), specify the `--pthread` option of the `g++` command.

8. If you are using arbitrary precision datatypes, include the reference version in your source code instead of the FPGA-optimized version provided with the Intel HLS Compiler. You can use the `__INTELFPGA_COMPILER__` macro to control which variant is included. For example, if you are using arbitrary precision integers, you can use the following macro code:

```c
#ifdef __INTELFPGA_COMPILER__
#include "HLS/ac_int.h"
#else
#include "ref/ac_int.h"
#endif
```

If you implement these steps, your `g++` command resembles the following example command:

```
g++ myFile.cpp -g -I"$(HLS_INSTALL_DIR)/include"
-L"$(HLS_INSTALL_DIR)/host/linux64/lib" -lhls_emul
-pthread -std=c++17
```

**Microsoft Visual C++**

The following instructions were tested with Microsoft Visual Studio 2017 Professional.

To compile your Intel HLS Compiler code with Microsoft Visual C++:

1. Initialize your environment with the Intel HLS Compiler initialization script:
   ```
   <quartus_installdir>/hls/init_hls.bat
   ```

2. Add the Intel HLS Compiler header files to the compiler command include path with the `/I` command option.
   The header files are in the `quartus_installdir\hls\include` directory.

3. Add the `/Zi` option to generate debug symbols when compiling.

4. Add the `/wd4068` option to suppress warnings because MSVC does not recognize the Intel HLS Compiler pragmas.

5. Add the HLS emulation library to the linker search path with the `/libpath` command option.
   The emulation library is in the `quartus_installdir\hls\host\windows64\lib` directory.

6. Add the `hls_emul` library to the linker command.

7. If you are using arbitrary precision datatypes, include the reference version instead of the FPGA-optimized version provided with the Intel HLS Compiler. You can use the `__INTELFPGA_COMPILER__` macro to control which version is included:

```c
#ifdef __INTELFPGA_COMPILER__
#include "HLS/ac_int.h"
#else
#include "ref/ac_int.h"
#endif
```
Your Microsoft Visual C++ compiler command should resemble the following example command:

```
cl myFile.cpp /I "%HLS_INSTALL_DIR%\include" /nologo /EHsc /wd4068 /MD
    /std:c++17 /Zi
    /link /libpath:%HLS_INSTALL_DIR%\host\windows64\lib
    hls_emul.lib
```

### 2.3.1. Compiling Your Testbench and Component Separately

In some cases, you might want to decouple your testbench development from your component development. For example, you might want to change your testbench code without recompiling your HLS component.

You can compile your testbench and component separately in one of the following ways:

- Compile your testbench with the Intel HLS Compiler and then compile your component code with the Intel HLS Compiler.
- Compile your testbench with GCC or Microsoft Visual Studio, and then compile your component code with the Intel HLS Compiler.

In either case, the build flow is similar:

1. **Use the `i++ -c` command to compile your component code to an object file.**
   
   For example, `i++ -c component.cpp -o component.o -march=<target_platform>`

2. **Use any compiler supported by the Intel HLS Compiler to compile your testbench to an object file (or files):**
   
   - Intel HLS Compiler: `i++ -c`
   - GCC: `g++ -c`
   - MSVC: `cl /c`

   For example, `g++ -c testbench.cpp -o testbench.o`

3. **Link the component and testbench object files with the `i++`.**
   
   For example, `i++ testbench.o component.o -o prog -march=<target_platform>`.

For an example of compiling your testbench and component separately, review the following tutorial:

```
<quartus_installdir>/hls/examples/tutorials/usability/compiler_interoperability
```

### 2.4. Intel HLS Compiler Hardware Model

The Intel HLS Compiler attempts to pipeline functions as much as possible. Different stages of the pipeline might have multiple operations performed in parallel.

The following figure shows an example of the pipeline architecture generated by the Intel HLS Compiler. The numbered operations on the right side represent the pipeline implementation of the C++ code on the left side of the figure. Each box in the right side of the figure is an operation in the pipeline.
With a pipelined approach, multiple invocations of the component can be simultaneously active. For example, the earlier figure shows that the first invocation of the component can be returning a result at the same time the fourth invocation of the component is called.

One invocation of a component advances to its next stage in the pipeline only after all of the operations of its current stage are complete.

Some operations can stall the pipeline. A common example of operations that can stall a pipeline is a variable latency operation like a memory load or store operation. To support pipeline stalls, the Intel HLS Compiler propagates ready and valid signals through the pipeline to all operations that have a variable latency.

For operations that have a fixed latency, the Intel HLS Compiler can statically schedule the interaction between the operations and ready signals are not needed between the stages with fixed latency operations. In these cases, the compiler optimizes the pipeline to statically schedule the operations, which significantly reduces the logic required to implement the pipeline.
3. C Language and Library Support

3.1. Supported C and C++ Subset for Component Synthesis

Some common software patterns cannot be physically realized in FPGA digital logic, which results in limitations in the coding style that the Intel HLS Compiler can support.

The compiler cannot generate RTL for the following C++ constructs:

- Dynamic memory allocation.
- Virtual functions.
- Function pointers
- C++ or C library functions, except the supported math functions explicitly mentioned in Supported Math Functions on page 204.
- Non-static class functions.
- Template functions without an explicit specialization.

In general, the compiler can generate RTL for functions that can be statically linked such as static class methods and "regular" functions. HLS component functions can include classes, structs, functions, templates, and pointers.

In addition, a component or task function cannot contain an irreducible loop. That is, loops in component and task functions must have only one entry point into the loop.

Important: These synthesis limitations do not apply to testbench code.

3.2. C and C++ Libraries

The Intel High Level Synthesis (HLS) Compiler provides a number of header files to provide FPGA implementations of certain C and C++ functions.

Table 5. Intel HLS Compiler Pro Edition Header Files Summary

<table>
<thead>
<tr>
<th>HLS Header File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLS/hls.h</td>
<td>Required for component identification and component parameter interfaces.</td>
</tr>
<tr>
<td>HLS/math.h</td>
<td>Includes FPGA-specific definitions for the math functions from the math.h for your operating system.</td>
</tr>
<tr>
<td>HLS/extendedmath.h</td>
<td>Includes additional FPGA-specific definitions of math functions not in math.h.</td>
</tr>
<tr>
<td>HLS/ac_int.h</td>
<td>Provides FPGA-optimized arbitrary width integer support.</td>
</tr>
<tr>
<td>HLS/ac_fixed.h</td>
<td>Provides FPGA-optimized arbitrary precision fixed point support.</td>
</tr>
<tr>
<td>HLS/ac_fixed_math.h</td>
<td>Provides FPGA-optimized arbitrary precision fixed point math functions.</td>
</tr>
<tr>
<td>HLS/ac_complex.h</td>
<td>Provides FPGA-optimized complex number support.</td>
</tr>
</tbody>
</table>

continued...
<table>
<thead>
<tr>
<th>HLS Header File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLS/hls_float.h</td>
<td>Provides FPGA-optimized arbitrary-precision IEEE-754 compliant floating-point number support.</td>
</tr>
<tr>
<td>HLS/hls_float_math.h</td>
<td>Provides FPGA-optimized floating-point math functions.</td>
</tr>
<tr>
<td>HLS/math_dsp_control.h</td>
<td>Provides the ihc::math_dsp_control function to control the hardware implementation for supported data types and math functions at a local-scope.</td>
</tr>
<tr>
<td>HLS/stdio.h</td>
<td>Provides printf support for components so that printf statements work in x86 emulations, but are disabled in component when compiling to an FPGA architecture.</td>
</tr>
<tr>
<td>&lt;iostream&gt;</td>
<td>To use cout and cerr in your component, guard the statements with the HLS_SYNTHESIS macro.</td>
</tr>
</tbody>
</table>

**math.h**

To access functions in math.h from your component function, include the "HLS/math.h" file in your source code. The header ensures that the components call the hardware versions of the math functions.

For more information about supported math.h functions, see Supported Math Functions on page 204.

**stdio.h**

Synthesized component functions generally do not support C and C++ standard library functions such as FILE objects.

A component can call printf by including the header file HLS/stdio.h. This header changes the behavior of printf depending on the compilation target:

- For compilation that targets the x86-64 architecture (that is, -march=x86-64), the printf call behaves as normal.
- For compilation that targets the FPGA architecture (that is, -march="<FPGA_family_or_part_number>"), the compiler removes the printf call.

If you use printf in a component function without first including the #include "HLS/stdio.h" line in your code, you get an error message similar to the following error when you compile hardware to the FPGA architecture:

```
$ i++ -march="<FPGA_family_or_part_number>" --component dut test.cpp
Error: HLS gen_qsys FAILED.
See ./a.prj/dut.log for details.
```

You can use C and C++ standard library functions such as fopen and printf as normal in all testbench functions.

**iostream**

The C++ standard output streams (cout or cerr) output streams are not supported in RTL, but you can use them to debug your component in emulation. However, you must guard any cout or cerr statements with the HLS_SYNTHESIS macro.
This macro ensures that statements in a component work in x86 emulations (that is, -march=x86-64), but are disabled in the component when compiling it to an FPGA architecture (that is, -march="<FPGA_family_or_part_number>"). For example:

```c
#include "HLS/hls.h"
#include <iostream>

component int debug_component (int a){
#ifndef HLS_SYNTHESIS
    std::cout << "input value: " << a << std::endl;
#endif
    return a;
}
```

If you attempt to use `cout` or `cerr` in a component function without guarding the line in your code with the `HLS_SYNTHESIS` macro, you get an error message similar to the following error when you compile hardware to the FPGA architecture:

```
$ i++ -march="<FPGA_family_or_part_number>" run.cpp
run.cpp:5: Compiler Error: Cannot synthesize std::cout used inside of a component.
HLS Main Optimizer FAILED.
```

**Related Information**

Supported Math Functions on page 204

### 3.3. Templated and Overloaded Functions

You can use templating and overloading to create generalized function interfaces for your HLS components and HLS tasks. HLS components can be both templated and overloaded. HLS tasks can only be templated. You cannot overload an HLS task function.

**Related Information**

Task Functions on page 101

#### 3.3.1. Templated Functions

Using a templated function as an HLS component differs from using the templated function as an HLS task.

**Templated Functions as an HLS Component**

When you create a template function, you must declare the instantiation of the function to synthesize into hardware.

For example, a templated `multadd` function might be useful in a system.

```c
template <typename T, int MULT>
T multadd (T a, T b) {
    return MULT * (a + b);
}
```

To synthesize a version of this function into a component, you must declare the instantiation that you want to synthesize:

```c
template component int multadd<int, 5>(int a, int b);
```
This declaration combined with the earlier template definition marks the int variant with MULT=5 of the multiadd function to be generated into a component. This component can now be invoked from the testbench.

**Templated Functions as an HLS Task**

If you want to use the function as a task in a system of tasks, use the ihc::launch and ihc::collect calls as shown in the following example:

```c
component void foo () {
    int a, b;
    ihc::launch<multadd<int, 5>> (a, b);
    int res = ihc::collect<multadd<int, 5>>();
}
```

### 3.3.2. Overloaded Functions

HLS component functions can be overloaded, but HLS task functions cannot because the ihc::launch and ihc::collect calls cannot distinguish between overloaded variants of a task function.

To overload a component function, define multiple variants of the function.

For example:

```c
component int mult (int a, int b) {
    return a * b;
}
```

```c
component float mult (float a, float b) {
    return a * b;
}
```

### 3.3.3. Function Name Mapping

The Intel HLS Compiler always generates unique function names to avoid name collisions that might occur for overloaded and templated functions.

A mapping of the full function declaration to the synthesized function name is provided in the summary page of the High-Level Design Reports (report.html). The synthesized function name is used for all the other reports such as the loops report and area analysis.

The following example shows an example of this table in the report:
3.4. Intel HLS Compiler Pro Edition Compiler-Defined Preprocessor Macros

The Intel HLS Compiler Pro Edition has built-in macros that you can use to customize your code to create flow-dependent behaviors.

Table 6. Macro Definition for __INTELFPGA_COMPILER__

<table>
<thead>
<tr>
<th>Tool Invocation</th>
<th><strong>INTELFPGA_COMPILER</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>g++ or cl</td>
<td>Undefined</td>
</tr>
<tr>
<td>i++ -march=x86-64</td>
<td>2120</td>
</tr>
<tr>
<td>i++ -march=&quot;&lt;FPGA_family_or_part_number&gt;&quot;</td>
<td>2120</td>
</tr>
</tbody>
</table>

Table 7. Macro Definition for HLS_SYNTHESIS

<table>
<thead>
<tr>
<th>Tool Invocation</th>
<th>HLS_SYNTHESIS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Testbench Code</td>
</tr>
<tr>
<td>g++ or cl</td>
<td>Undefined</td>
</tr>
<tr>
<td>i++ -march=x86-64</td>
<td>Undefined</td>
</tr>
<tr>
<td>i++ -march=&quot;&lt;FPGA_family_or_part_number&gt;&quot;</td>
<td>Undefined</td>
</tr>
</tbody>
</table>
4. Component Interfaces

The Intel HLS Compiler generates a component interface for integrating your RTL component into a larger system. A component has two basic interface types: the component invocation interface and the parameter interface.

The **component invocation interface** is common to all HLS components and contains the return data (for non-`void` functions) and handshake signals for invoking the component, and for receiving results back when the component finishes executing.

Use the **parameter interface** to transfer data in and out of your component function. The parameter interface for your component is based on the parameters that you define in your component function signature and global variables (including global streams) that your component accesses.

4.1. Component Invocation Interface

For each function that you label as a component, the Intel HLS Compiler creates a corresponding RTL module. This RTL module must have top-level ports, or interfaces, that allow your overall system to interact with your HLS component.

By default, the RTL module for an HLS component includes the following interfaces and data:

- A call interface that consists of `start` and `busy` signals. The call interface is sometimes referred to as the **do stream**.
- A return interface that consists of `done` and `stall` signals. The return interface is sometimes referred to as the **return stream**.
- Return data if the component function has a return type that is not `void`

The following diagram shows a component with the default call and return interfaces:

```
  component
  +--------+    +--------+
  |  start |    |  busy  |
  +--------+    +--------+
    |        |    |        |
    +--------+    +--------+
        |        |    |        |
        |        |    |        |
        |        +--------+
        |                |  done
        |                |  stall
```

Alternatively, by declaring your component as an `hls_avalon_agent_component` component, your component can have signals registered in the component agent memory map instead. In an `hls_avalon_agent_component` component, the `start`, `done`, and `returndata` signals appear in the component control and status registers (CSR) instead of as conduits outside of the component.

For a comparison of the invocation interfaces, see Interface Definition Example: Component Invocation Interface Control Attributes on page 23.
For an example of a component interface with scalar and pointer arguments, see Figure 2 on page 24.

**Interfaces and Generated RTL**

Your component function parameters generate different RTL depending on their type. For details see the following sections:

- Scalar Parameters on page 22
- Pointer and Reference Parameters on page 23

You can also explicitly declare Avalon Streaming interfaces (using `stream_in<>` and `stream_out<>` classes) and Avalon Memory-Mapped Host (using `mm_host<>` classes) interfaces on component interfaces. For details see the following sections:

- Avalon Streaming Interfaces on page 25
- Avalon Memory-Mapped Host Interfaces on page 32

**Component Invocation Interface Control Attributes**

You can indicate the control signals that correspond to the actions of calling your component by using one of the component invocation interface attributes.

Unless a component parameter is marked stable (with the `hls_stable_argument` attribute), the component parameter inputs are synchronized according to this component invocation protocol.

**Table 8. Intel HLS Compiler Component Invocation Interface Control Attribute Summary**

<table>
<thead>
<tr>
<th>Control Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>hls_avalon_streaming_component</code></td>
<td>This is the default component invocation interface.</td>
</tr>
<tr>
<td></td>
<td>The component uses <code>start</code>, <code>busy</code>, <code>stall</code>, and <code>done</code> signals for handshaking.</td>
</tr>
<tr>
<td><code>hls_avalon_agent_component</code></td>
<td>The <code>start</code>, <code>done</code>, and <code>returndata</code> (if applicable) signals appear in the component CSR instead of as conduits outside of the signal.</td>
</tr>
<tr>
<td><code>hls_always_run_component</code></td>
<td>The <code>start</code> signal is tied to 1 internally in the component. There is no <code>done</code> signal output.</td>
</tr>
<tr>
<td><code>hls_stall_free_return</code></td>
<td>The <code>stall</code> signal is removed by internally setting it to 0. Use this control attribute only if the downstream component never stalls.</td>
</tr>
</tbody>
</table>

**Related Information**

- Control and Status Register (CSR) Agent on page 43
- Stable Component Parameters on page 48

**4.1.1. Scalar Parameters**

Each scalar parameter in your component results in an input conduit that is synchronized with the component `start` and `busy` signals.

The inputs are read into the component when the external system pulls the `start` signal high and the component keeps the `busy` signal low.
If your component has a return value, an output conduit that is synchronized to your component done signal is generated.

For an example of how to specify a scalar parameter and how it is read in by a component, see the a argument in Figure 2 on page 24 and Figure 3 on page 25.

### 4.1.2. Pointer and Reference Parameters

Each pointer or reference parameter of a component results in an input conduit for the address. The input conduit is synchronized with the component start and busy signals. In addition to this input conduit, all pointers share a single Avalon Memory-Mapped (MM) host interface that the component uses to access system memory.

You can customize these pointer interfaces using the `mm_host<>` class.

*Note:* Explicitly-declared Avalon Memory-Mapped Host interfaces and Avalon Streaming interfaces are passed by reference.

For details about Avalon (MM) Host interfaces, see Avalon Memory-Mapped Host Interfaces on page 32.

### 4.1.3. Interface Definition Example: Component Invocation Interface

**Control Attributes**

This example compares two simple components. One component is implemented with simple conduits as its signal interface, while the other component is implemented as an `hls_avalon_agent_component`.

Consider the following code example for a simple component:

```cpp
component float myComponent(float a, float b) {
  return a+b;
}
```

This code example results in a component with the following signals:

```
start
busy
[a[31:0]]
[b[31:0]]
returndata[31:0]
```

You can also implement this component as an `hls_avalon_agent_component` component with the control signals residing in a register map:

```cpp
hls_avalon_agent_component
component float myComponent(
  hls_avalon_agent_register_arg float a,
  hls_avalon_agent_register_arg float b) {
  return a+b;
}
```
This code results in a component with the following signals:

```
component int dut(int a, int* b, int i) {
    return a*b[i];
}
```

### 4.1.4. Interface Definition Example: Component with Both Scalar and Pointer Arguments

The following design example illustrates the interactions between a component’s interfaces and signals, and the waveform of the corresponding RTL module.

**Component**

```
component int dut(int a, int* b, int i) {
    return a*b[i];
}
```

**Figure 2.** Block Diagram of the Interfaces and Signals for the Component dut
Waveform Diagram of the Signals for the Component dut

This diagram shows that the Avalon-MM read signal reads from a memory interface that has a read latency of one cycle and is non-blocking.

If the `dut` component raises the `busy` signal, the caller needs to keep the `start` signal high and continue asserting the input arguments. Similarly, if the component downstream of `dut` raises the `stall` signal, then `dut` holds the `done` signal high until the `stall` signal is de-asserted.

4.2. Avalon Streaming Interfaces

A component can have input and output streams that conform to the Avalon Streaming (ST) interface specifications. These input and output streams are represented by passing references to `ihc::stream_in<>` and `ihc::stream_out<>` objects as function arguments to the component.

When you use an Avalon ST interface, you can serialize the data over several clock cycles. That is, one component invocation can read from a stream multiple times.

You cannot derive new classes from the stream classes or encapsulate them in other formats such as structs or arrays. However, you may use references to instances of these classes as references inside other classes, meaning that you can create a class that has a reference to a stream object as a data member.

A component can have multiple read sites for a stream. Similarly, a component can have multiple write sites for a stream. However, for best component performance try to restrict each input stream in your design to a single read site and each output stream to a single write site.

**Note:** Within the component, there is no guarantee on the order of execution of different streams unless a data dependency exists between streams.

For more information about streaming interfaces, refer to "Avalon Streaming Interfaces" in *Avalon Interface Specifications*.

**Restriction:** The Intel HLS Compiler does not support the Avalon ST channel or error signals.
## Streaming Input Interfaces

Table 9. Intel HLS Compiler Pro Edition Streaming Input Interface Template Summary

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::stream_in</td>
<td>Streaming input interface to the component.</td>
</tr>
<tr>
<td>ihc::buffer</td>
<td>Specifies the capacity (in words) of the FIFO buffer on the input data that associates with the stream.</td>
</tr>
<tr>
<td>ihc::readyLatency</td>
<td>Specifies the number of cycles between when the ready signal is deasserted and when the input stream can no longer accept new inputs.</td>
</tr>
<tr>
<td>ihc::bitsPerSymbol</td>
<td>Describes how the data is broken into symbols on the data bus.</td>
</tr>
<tr>
<td>ihc::firstSymbolInHighOrderBits</td>
<td>Specifies whether the data symbols in the stream are in big endian order.</td>
</tr>
<tr>
<td>ihc::usesPackets</td>
<td>Exposes the startofpacket and endofpacket sideband signals on the stream interface.</td>
</tr>
<tr>
<td>ihc::usesEmpty</td>
<td>Exposes the empty out-of-band signal on the stream interface.</td>
</tr>
<tr>
<td>ihc::usesValid</td>
<td>Controls whether a valid signal is present on the stream interface.</td>
</tr>
</tbody>
</table>

Table 10. Intel HLS Compiler Pro Edition Streaming Input Interface stream_in Function APIs

<table>
<thead>
<tr>
<th>Function API</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T read()</td>
<td>Blocking read call to be used from within the component.</td>
</tr>
<tr>
<td>T read(bool&amp; sop, bool&amp; eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Blocking read with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>T read(bool&amp; sop, bool&amp; eop, int&amp; empty)</td>
<td>Available only if usesPackets&lt;true&gt; and usesEmpty&lt;true&gt; are set. Blocking read with out-of-band startofpacket, endofpacket, and empty signals.</td>
</tr>
<tr>
<td>T tryRead(bool &amp;success)</td>
<td>Non-blocking read call to be used from within the component. The success bool is set to true if the read was valid. That is, the Avalon-ST valid signal was high when the component tried to read from the stream. The emulation model of tryRead() is not cycle-accurate, so the behavior of tryRead() might differ between emulation and simulation.</td>
</tr>
<tr>
<td>T tryRead(bool&amp; success, bool&amp; sop, bool&amp; eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Non-blocking read with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>T tryRead(bool&amp; success, bool&amp; sop, bool&amp; eop, int&amp; empty)</td>
<td>Available only if usesPackets&lt;true&gt; and usesEmpty&lt;true&gt; are set. Non-blocking read with out-of-band startofpacket, endofpacket, and empty signals.</td>
</tr>
</tbody>
</table>

continued...
Function API | Description
---|---
void write(T data) | Blocking write call to be used from the testbench to populate the FIFO to be sent to the component.
void write(T data, bool sop, bool eop) | Available only if usesPackets<true> is set. Blocking write call with out-of-band startofpacket and endofpacket signals.
void write(T data, bool sop, bool eop, int empty) | Available only if usesPackets<true> and usesEmpty<true> are set. Blocking write call with out-of-band startofpacket, endofpacket, and empty signals.

## Streaming Output Interfaces

### Table 11. Intel HLS Compiler Pro Edition Streaming Output Interface Template Summary

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::stream_out</td>
<td>Streaming output interface from the component.</td>
</tr>
<tr>
<td>ihc::readylatency</td>
<td>Specifies the number of cycles between when the ready signal is deasserted and when the input stream can no longer accept new inputs.</td>
</tr>
<tr>
<td>ihc::bitsPerSymbol</td>
<td>Describes how the data is broken into symbols on the data bus.</td>
</tr>
<tr>
<td>ihc::firstSymbolInHighOrderBits</td>
<td>Specifies whether the data symbols in the stream are in big endian order.</td>
</tr>
<tr>
<td>ihc::usesPackets</td>
<td>Exposes the startofpacket and endofpacket sideband signals on the stream interface.</td>
</tr>
<tr>
<td>ihc::usesEmpty</td>
<td>Exposes the empty out-of-band signal on the stream interface.</td>
</tr>
<tr>
<td>ihc::usesReady</td>
<td>Controls whether a ready signal is present.</td>
</tr>
</tbody>
</table>

### Table 12. Intel HLS Compiler Pro Edition Streaming Output Interface stream_out Function APIs

<table>
<thead>
<tr>
<th>Function API</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>void write(T data)</td>
<td>Blocking write call from the component</td>
</tr>
<tr>
<td>void write(T data, bool sop, bool eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Blocking write call with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>void write(T data, bool sop, bool eop, int empty)</td>
<td>Available only if usesPackets&lt;true&gt; and usesEmpty&lt;true&gt; are set. Blocking write call with out-of-band startofpacket, endofpacket, and empty signals.</td>
</tr>
<tr>
<td>bool tryWrite(T data)</td>
<td>Non-blocking write call from the component. The return value represents whether the write was successful.</td>
</tr>
<tr>
<td>bool tryWrite(T data, bool sop, bool eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Non-blocking write with out-of-band startofpacket and endofpacket signals. The return value represents whether the write was successful. That is, the downstream interface was pulling the ready signal high while the HLS component tried to write to the stream.</td>
</tr>
</tbody>
</table>

**continued...**
### Function API

<table>
<thead>
<tr>
<th>Function API</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>bool tryWrite(T data, bool sop, bool eop, int empty)</td>
<td>Available only if usesPackets&lt;true&gt; and usesEmpty&lt;true&gt; are set. Non-blocking write with out-of-band startofpacket, endofpacket, and empty signals. The return value represents whether the write was successful.</td>
</tr>
<tr>
<td>T read()</td>
<td>Blocking read call to be used from the testbench to read back the data from the component</td>
</tr>
<tr>
<td>T read(bool &amp;sop, bool &amp;eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Blocking read call to be used from the testbench to read back the data from the component with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>T read(bool &amp;sop, bool &amp;eop, int &amp;empty)</td>
<td>Available only if usesPackets&lt;true&gt; and usesEmpty&lt;true&gt; are set. Blocking read call to be used from the testbench to read back the data from the component with out-of-band startofpacket, endofpacket, and empty signals.</td>
</tr>
</tbody>
</table>

### Related Information

Avalon Interface Specifications

### 4.3. Pipes

Using external memory to move data in and out of static-object libraries or task functions can constrain the performance of your design. Pipes provide a mechanism for passing data with high efficiency and low latency by using on-device FIFO buffers to communicate.

Pipes are similar to streams, but are simpler. You can use pipes in a component or to allow tasks to communicate with each other, but pipes support only ready/valid and data signals, while streams support additional sideband signals such as startofpacket and endofpacket.

Unlike streams, you can declare an array of pipes and you can iterate over the array to write or read many pipes in a design.

Data written to a pipe remains in the pipe until it is read or until the component is reset.

The memory model of pipes allows you to use them to send and receive data from task functions in running functions in a static-object library or a running task function.

### Pipe Properties

Pipes have the following key properties:

- **FIFO ordering**: Data is accessible (readable) only in FIFO order. There is no concept of a memory address or pointer in a pipe, which means random data access is not possible with pipes.

- **Capacity**: Pipes have a capacity. That is, a fixed amount of data can be written to an initially empty pipe before needing to read anything from it to make room for more data.
A full pipe applies back pressure to the write site.

**Pipe Accessors**

Data is written to a pipe through an API that commits a single word of the pipe data type, and that word is later returned by an API that reads data from the pipe.

The API accessing the pipe can be a blocking or a nonblocking type:

- **Blocking API calls**
  - Blocking write API calls wait until the pipe has enough capacity to commit data.
  - Blocking read API calls wait until the pipe contains data to be read.

- **Nonblocking API calls**
  - Nonblocking API calls execute immediately and return a status that indicates whether the operation was successful.

You can mix blocking and nonblocking accesses to the same pipe. For example you can write data to a pipe with a blocking pipe `write()` call and read it from the other end using a non-blocking `read()` call, and vice versa.

**Data Persistence in Pipes**

Data written to a pipe by a component remains in the pipe until it is read or until the component is reset.

Data written to a pipe by a task remains in the pipe until another task reads from the pipe or the component containing the tasks is reset.

The sequence of data in a pipe always follows FIFO ordering.

Data in pipes does not persist across FPGA device resets or reprogramming.

**Restrictions on using Pipes**

Using pipes comes with the following restrictions:

- **Multiple pipe call sites**
  - A component or task function can read from the same pipe multiple times, but multiple component or task functions cannot read from the same pipe. Similarly, a component or task function can write to the same pipe multiple times, but multiple component or task functions cannot write to the same pipe.

- **Feedback and feed-forward pipes**
  - A component or task function should use separate pipes for pipe reads and pipe writes. Writing to and reading from the same pipe within the same function can lead to poor performance.
Pipe accesses in loops

Do not use nonblocking pipes if you use a loop structure that waits for data from the pipe. That is, avoid the following code pattern for nonblocking pipe accesses:

```cpp
bool success = false;
while (!success) {
    my_pipe::write(rd_src_buf[i], success);
    // can also be a nonblocking read
}
```

Use a blocking access with this code pattern instead because a blocking access in this code pattern is more efficient in hardware than a nonblocking access in this code pattern.

Related Information

- Static-Object Libraries on page 112
- Systems of Tasks on page 100

4.3.1. The pipe Class and Its Use

The pipe class exposes static methods for writing a data word to a pipe and reading a data word from a pipe. The reads and writes can be blocking or nonblocking, with the form chosen based on the overload resolution.

The pipe API is equivalent to the following class declaration:

```cpp
template <class name,
         class dataT,
         size_t min_capacity = 0>
class pipe {
public:
    // Blocking
    static dataT read();
    static void write(dataT data);
    // Non-blocking
    static dataT read(bool &success);
    static void write(dataT data, bool &success);
}
```

Where the template parameters are defined as follows:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>name</td>
<td>The type that is used to create a unique identifier for the pipe. It is typically a user-defined class, in a user namespace. Forward declaration of the type is enough, and the type need not be defined.</td>
</tr>
<tr>
<td>dataT</td>
<td>The data type of the packet contained within a pipe. This is the data type that is read during a successful pipe <code>read()</code> operation, or written during a successful pipe <code>write()</code> operation. The type must have a standard layout and be trivially copyable.</td>
</tr>
<tr>
<td>min_capacity</td>
<td>The minimum number of words (in units of <code>dataT</code>) that the pipe must be able to store without any being read out. The compiler might create a pipe with a larger capacity due to performance considerations.</td>
</tr>
</tbody>
</table>
Example Using Pipes

The following code example shows a header file that you can use in a source code library or use in component or task functions.

```cpp
#include "HLS/hls.h"

template<unsigned ID, class T, unsigned pipe_capacity> class TaskSystem {
private:
    template<unsigned SystemID> class InputPipeID {};
    template<unsigned SystemID> class TaskPipeID {};
    template<unsigned SystemID> class OutputPipeID {};

public:
    using input_pipe  = ihc::pipe<class InputPipeID<ID>, T, pipe_capacity>;
    using output_pipe = ihc::pipe<class OutputPipeID<ID>, T, pipe_capacity>;
    using task_pipe   = ihc::pipe<class TaskPipeID<ID>, T, pipe_capacity>;

    static void first_task(unsigned N) {
        T data;
        for(unsigned i=0; i<N; ++i) {
            data = input_pipe::read();
            task_pipe::write(data);
        }
    }

    static void second_task(unsigned N) {
        T data;
        for(unsigned i=0; i<N; ++i) {
            data = task_pipe::read();
            output_pipe::write(data);
        }
    }
};
```

With this header file, `first_task` and `second_task` can be called from separate task functions to achieve concurrency.

Assuming the header file in a file called `task_system.h`, you can create a system of tasks component like the following example:

```cpp
#include "HLS/hls.h"
#include <iostream>

using MySystem = TaskSystem<ID, int, CAPACITY>;

int main() {
    ihc::launch<MySystem::first_task>(CAPACITY);
    ihc::launch<MySystem::second_task>(CAPACITY);

    for(int i = 0; i < CAPACITY; ++i) {
        std::cout << "input: " << i << "\n";
        MySystem::input_pipe::write(i);
    }

    for(int i = 0; i < CAPACITY; ++i) {
        int data = MySystem::output_pipe::read();
        std::cout << "output: " << data << "\n";
    }

    return 0;
}
```
Example of an Array of Pipes

The following code example implements an array of pipes using templates, and includes functions to write to such an array.

```cpp
#include "HLS/hls.h"

// PipeArray
template <class ArrayID, typename T, unsigned pipeCapacity, unsigned arraySize>
class PipeArray {
private:
  template <unsigned idx> struct StructIndex;
  template <unsigned idx> struct VerifyIndex {
    static_assert(idx < arraySize, "Index out of bounds");
    using VerifiedPipe = ihc::pipe<StructIndex<idx>, T, pipeCapacity>;
  };
public:
  template <unsigned idx>
  using pipe_at = typename VerifyIndex<idx>::VerifiedPipe;
  static void write_to_pipes(T *values);
};

// Write Unroller
template <class ArrayID, typename T, unsigned pipeCapacity, unsigned arraySize,
unsigned idx>
struct WriteUnroller {
  using my_array = PipeArray<ArrayID, T, pipeCapacity, arraySize>;
  static void write_to_pipes_impl(T *values) {
    my_array::template pipe_at<idx>::write(values[idx]);
    WriteUnroller<ArrayID, T, pipeCapacity, arraySize, idx + 1>::write_to_pipes_impl(values);
  }
};

// Write function
template <class ArrayID, typename T, unsigned pipeCapacity, unsigned arraySize>
void PipeArray<ArrayID, T, pipeCapacity, arraySize>::write_to_pipes(T *values) {
  WriteUnroller<ArrayID, T, pipeCapacity, arraySize, 0>::write_to_pipes_impl(values);
}
```

The function `PipeArray::write_to_pipes` takes an array of values to be written, and calls `WriteUnroller::write_to_pipes_impl`, which uses recursive templating to write to each pipe in the array. Reading from the array of pipes would have a similar implementation.

4.4. Avalon Memory-Mapped Host Interfaces

A component can interface with an external memory over an Avalon Memory-Mapped (MM) Host interface. You can specify the Avalon MM Host interface implicitly using a function pointer argument or reference argument, or explicitly using the `mm_host<>` class defined in the "HLS/hls.h" header file. Describe a customized Avalon MM Host interface in your code by including a reference to an `mm_host<>` object in your component function signature.
Each `mm_host` argument of a component results in an input conduit for the address. That input conduit is associated with the component start and busy signals. In addition to this input conduit, a unique Avalon MM Host interface is created for each address space. Host interfaces that share the same address space are arbitrated on the same interface.

For more information about Avalon MM Host interfaces, refer to "Avalon Memory-Mapped Interfaces" in *Avalon Interface Specifications*.

### Table 14. Intel HLS Compiler Pro Edition Memory-Mapped Interfaces Summary

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ihc::mm_host</code></td>
<td>The underlying pointer type.</td>
</tr>
<tr>
<td><code>ihc::dwidth</code></td>
<td>The width of the memory-mapped data bus in bits.</td>
</tr>
<tr>
<td><code>ihc::awidth</code></td>
<td>The width of the memory-mapped address bus in bits.</td>
</tr>
<tr>
<td><code>ihc::aspace</code></td>
<td>The address space of the interface that associates with the host.</td>
</tr>
<tr>
<td><code>ihc::latency</code></td>
<td>The guaranteed latency from when a read command exits the component when the external memory returns valid read data.</td>
</tr>
<tr>
<td><code>ihc::maxburst</code></td>
<td>The maximum number of data transfers that can associate with a read or write transaction.</td>
</tr>
<tr>
<td><code>ihc::align</code></td>
<td>The alignment of the base pointer address in bytes.</td>
</tr>
<tr>
<td><code>ihc::readwrite_mode</code></td>
<td>The port direction of the interface.</td>
</tr>
<tr>
<td><code>ihc::waitrequest</code></td>
<td>Adds the <code>waitrequest</code> signal that is asserted by the agent when it is unable to respond to a read or write request.</td>
</tr>
<tr>
<td><code>getInterfaceAtIndex</code></td>
<td>This testbench function is used to index into an <code>mm_host</code> object.</td>
</tr>
</tbody>
</table>

**Related Information**

*Avalon Interface Specifications*

### 4.4.1. Memory-Mapped Host Testbench Constructor

For components that use an instance of the Avalon Memory-Mapped (MM) Host class (`mm_host<>`) to describe their memory interfaces, you must create an `mm_host<>` object in the testbench for each `mm_host` argument.

To create an `mm_host<>` object, add the following constructor in your code:

```cpp
ihc::mm_host<int, ... > mm(void* ptr, int size, bool use_socket=false);
```
where the constructor arguments are as follows:

- **ptr** is the underlying pointer to the memory in the testbench
- **size** is the total size of the buffer in bytes
- **use_socket** is the option you use to override the copying of the memory buffer and have all the memory accesses pass back to the testbench memory

By default, the Intel HLS Compiler copies the memory buffer over to the simulator and then copies it back after the component has run. In some cases, such as pointer-chasing in linked lists, copying the memory buffer back and forth is undesirable. You can override this behavior by setting `use_socket` to `true`.

*Note:* When you set `use_socket` to `true`, only Avalon MM Host interfaces with 64-bit wide addresses are supported. In addition, setting this option increases the run time of the simulation.

### 4.4.2. Implicit and Explicit Examples of Describing a Memory Interface

Optimize component code that describes a memory interface by specifying an explicit `mm_host` object.

**Implicit Example**

The following code example arbitrates the load and store instructions from both pointer dereferences to a single interface on the component's top-level module. This interface will have a data bus width of 64 bits, an address width of 64 bits, and a fixed latency of 1.

```c
#include "HLS/hls.h"
component void dut(int *ptr1, int *ptr2) {
  *ptr1 += *ptr2;
  *ptr2 += ptr1[1];
}

int main(void) {
  int x[2] = {0, 1};
  int y = 2;
  dut(x, &y);
  return 0;
}
```

**Explicit Example**

This example demonstrates how to optimize the previous code snippet for a specific memory interface using the explicit `mm_host` class. The `mm_host` class has a defined template, and it has the following characteristics:

- Each interface is given a unique ID that infers two independent interfaces and reduces the amount of arbitration within the component.
- The data bus width is larger than the default width of 64 bits.
- The address bus width is smaller than the default width of 64 bits.
- The interfaces have a fixed latency of 2.

By defining these characteristics, you state that your system returns valid read data after exactly two clock cycles and that the interface never stalls for both reads and writes, but the system must be able to provide two different memories. A unique
physical Avalon MM host port (as specified by the \texttt{aspace} parameter) is expected to correspond to a unique physical memory. If you connect multiple Avalon MM Host interfaces with different physical Avalon MM host ports to the same physical memory, the Intel HLS Compiler cannot ensure functional correctness for any memory dependencies.

```c
#include "HLS/hls.h"

typedef ihc::mm_host<int, ihc::dwidth<256>,
           ihc::awidth<32>,
           ihc::aspace<1>,
           ihc::latency<2>> Host1;

typedef ihc::mm_host<int, ihc::dwidth<256>,
           ihc::awidth<32>,
           ihc::aspace<4>,
           ihc::latency<2>> Host2;

cOMPONENT void dut(Host1 &mm1, Host2 &mm2) {
  *mm1 += *mm2;
  *mm2 += mm1[1];
}

cOMPONENT void main(void) {
  int x[2] = {0, 1};
  int y = 2;
  Host1 mm_x(x, 2*sizeof(int), false);
  Host2 mm_y(&y, sizeof(int), false);
  dut(mm_x, mm_y);
  return 0;
}
```

### 4.4.3. Avalon Memory-Mapped Host Interfaces and Load-Store Units

When your component uses one or more Avalon Memory-Mapped (MM) Host interfaces, the Intel HLS Compiler inserts load-store units (LSUs) in the datapath between the interface and the rest of your component datapath. The type of LSU inserted depends on the inferred memory access pattern and other memory attributes.

The Intel HLS Compiler also tries to minimize the number of LSUs created by coalescing multiple load/store operations into wider load/store operations. Multiple LSUs can share a memory interface.

Typically, the Intel HLS Compiler creates burst-coalesced LSUs for variable-latency MM Host interfaces and pipelined LSUs for fixed-latency MM Host interfaces.

For details about the types of the LSUs and when the Intel HLS Compiler typically instantiates them, see Load-Store Unit Types on page 36 and Memory-Access Coalescing and Load-Store Units on page 40.

If your design contains one or more variable-latency Avalon MM Host interfaces (for example, if you interface with off-chip memory), you can control the LSU type to improve the performance and resource utilization of your design.

Use the high-level design reports to determine what types of LSUs your component has, and then you can apply these LSU controls as needed to achieve the component performance that you want.
Table 15. Intel HLS Compiler Pro Edition Load-Store Unit Control Summary

<table>
<thead>
<tr>
<th>Template Object/Parameter/Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::lsu</td>
<td>The underlying LSU class template object</td>
</tr>
<tr>
<td>ihc::style</td>
<td>Specifies the type of load-store unit.</td>
</tr>
<tr>
<td>ihc::static_coalescing</td>
<td>Explicitly allows or prevents static coalescing of a load/store operation</td>
</tr>
<tr>
<td></td>
<td>with other load/store operations.</td>
</tr>
<tr>
<td>load</td>
<td>Loads data from memory into the LSU.</td>
</tr>
<tr>
<td>store</td>
<td>Stores data from the LSU into memory.</td>
</tr>
</tbody>
</table>

Related Information
Control LSUs For Your Variable-Latency MM Host Interfaces

4.4.3.1. Load-Store Unit Types

The Intel HLS Compiler determines the types of load-store units (LSUs) to instantiate and whether to coalesce memory accesses based on the memory access pattern that the compiler infers.

The Intel HLS Compiler instantiates the following types of LSUs:

- **Burst-coalesced LSUs**: The Intel HLS Compiler typically instantiates burst-coalesced LSUs for accessing variable-latency Avalon MM Host interfaces.

- **Nonaligned burst-coalesced LSUs**: The Intel HLS Compiler typically instantiates burst-coalesced LSUs for accessing variable-latency Avalon MM Host interfaces.

- **Pipelined LSUs**: The Intel HLS Compiler typically instantiates pipelined LSUs for accessing fixed-latency Avalon MM Host interfaces.

- **Never-stall pipelined LSUs**: The Intel HLS Compiler typically instantiates pipelined LSUs for accessing fixed-latency Avalon MM Host interfaces.

Click LSUs in the System Viewer (in the High-Level Design Reports) to see which types of LSU the compiler instantiated for your component.
Burst-Coalesced Load-Store Units

By default, the compiler infers burst-coalesced load-store units (LSUs) for any variable-latency Avalon MM Host interface.

A burst-coalesced LSU dynamically buffers contiguous memory requests until the largest possible burst can be made or until the LSU receives no new requests for a given period of time. The largest possible burst is defined by the `ihc::maxburst` parameter. For noncontiguous memory requests, a burst-coalesced LSU flushes the buffer between requests.

Burst-coalesced LSUs provide efficient, variable-latency access to memories outside of your component. However, they require a considerable amount of FPGA resources.

The following code example results in the Intel HLS Compiler instantiating two burst-coalesced LSUs by default (because of the variable-latency Avalon MM Host interface):

```c
#include "HLS/hls.h"

component void
burst_coalesced(ihc::mm_host<int, ihc::dwidth<64>, ihc::awidth<32>,
                 ihc::aspace<1>, ihc::latency<0>> &in,
                 ihc::mm_host<int, ihc::dwidth<64>, ihc::awidth<32>,
                 ihc::aspace<2>, ihc::latency<0>> &out,
                 int i) {
    int value = in[i / 2]; // Burst-coalesced LSU
    out[i] = value; // Burst-coalesced LSU
}
```

Depending on the memory access pattern and other attributes, the compiler might modify a burst-coalesced LSU to be a nonaligned burst-coalesced LSU.
Nonaligned Burst-coalesced LSUs

When a burst-coalesced LSU can access a memory that is not aligned to the external memory word size, the Intel HLS Compiler creates a nonaligned burst-coalesced LSU. Nonaligned LSUs typically require more FPGA resources to implement than aligned LSUs. The throughput of a nonaligned LSU might be reduced if it receives many unaligned requests.

The following code example results in two nonaligned burst-coalesced LSUs:

```c
#include "HLS/hls.h"

struct State {
    int x;
    int y;
    int z;
};

cOMPONENT void static_coalescing(
    ihc::mm_host<State, ihc::dwidth<128>, ihc::awidth<32>,
    ihc::aspace<1>, ihc::latency<0>>& in,
    ihc::mm_host<State, ihc::dwidth<128>, ihc::awidth<32>,
    ihc::aspace<2>, ihc::latency<0>>& out,
    int i) {
    out[i] = in[i]; // Two Nonaligned Burst-coalesced LSUs
}
```

The figure that follows (Figure 5 on page 39) shows the external memory contents for the previous code example and the nonaligned burst-coalesced LSUs in the component pipeline.

The data type that is read and written is a 96-bit-wide struct. The external memory width is 128 bits. This difference between the read/write data width and the external memory width forces some of the memory requests to span two consecutive memory words.

A nonaligned burst-coalesced LSU can detect that discrepancy and serve such memory requests as needed while still buffering contiguous requests until the largest possible burst can be made.
Pipelined Load-Store Units

By default, the compiler infers pipelined load-store units (LSUs) for any fixed-latency Avalon MM Host interface and on-device memories.

In a pipelined LSU, requests are submitted when they are received and no buffering occurs. Pipelined LSUs are also used for accessing memories inside your component.

You can tell the compiler to instantiate pipelined LSUs for variable-latency MM Host interfaces. However, variable-latency interface access with pipelined LSUs might reduce throughput because pipelined LSUs do not combine sequential memory requests into bursts.

Memory accesses are pipelined, so multiple requests can be in flight at the same time.

The following code example results in the Intel HLS Compiler instantiating four pipelined LSUs:

```c
#include "HLS/hls.h"

component void pipelined(ihc::mm_host<int, ihc::dwidth<64>, ihc::awidth<32>, ihc::aspace<1>, ihc::latency<2>> &in,
                          ihc::mm_host<int, ihc::dwidth<64>, ihc::awidth<32>, ihc::aspace<1>, ihc::latency<2>> &out,
                          int gi, int li) {
    int lmem[1024];
    int res = in[gi]; // Pipelined LSU
    for (int i = 0; i < 4; i++) {
        lmem[li - i] = res; // Pipelined LSU
        res >>= 1;
    }
```
res = 0;
for (int i = 0; i < 4; i++) {
    res ^= lmem[li - i]; // Pipelined LSU
}
out[gi] = res; // Pipelined LSU

Never-Stall Pipelined LSUs

If a pipelined LSU is connected to a memory inside the component or to a fixed-latency MM Host interface without arbitration, a never-stall LSU is created because all accesses to the memory take a fixed number of cycles that are known to the compiler.

The following code example results in the Intel HLS Compiler instantiating three never-stall pipelined LSUs for accessing array lmem.

```c
#include "HLS/hls.h"

component void neverstall(ihc::mm_host<int, ihc::dwidth<128>, ihc::awidth<32>,
    ihc::aspace<1>, ihc::latency<0>> &in,
    ihc::mm_host<int, ihc::dwidth<128>, ihc::awidth<32>,
    ihc::aspace<1>, ihc::latency<0>> &out,
    int gi, int li) {
    int lmem[1024];
    for (int i = 0; i < 1024; i++)
        lmem[i] = in[i]; // Pipelined never-stall LSU
    out[gi] = lmem[li] ^ lmem[li + 1]; // Pipelined never-stall LSU
}
```

Related Information

- Intel HLS Compiler Pro Edition Load-Store Unit Control on page 194
- Burst Transfers in the Avalon Interface Specifications
- Intel HLS Compiler Pro Edition Memory-Mapped Interfaces on page 190

4.4.3.2. Memory-Access Coalescing and Load-Store Units

The Intel HLS Compiler sometimes coalesces multiple memory accesses into a wider memory access to save on the number of LSUs instantiated.

When the compiler coalesces the memory accesses, it is referred to as static coalescing because the coalescing occurs at compile time. This static coalescing contrasts with the dynamic coalescing done by a burst-coalesced LSU.

The compiler typically attempts static coalescing when it detects multiple memory operations that access consecutive locations in memory. This coalescing is usually beneficial because it reduces the number of LSUs that compete for a shared memory interface.

The compiler coalesces memory accesses only up to the width of the memory interface that is being accessed. For an external memory interface, the maximum width is predetermined by the properties of the external memory that you are accessing. For a component (internal) memory interface, the maximum width can be set by the compiler based on the memory geometry that the compiler creates. For more details about component memories, see Component Memories (Memory Attributes) on page 52.
For the following code example, the Intel HLS Compiler statically coalesces the four 4-byte-wide load operations into one 16-byte-wide load operation. A similar coalescing is done for the four store operations. Coalescing the load and store operations reduces the number of required accesses to the Avalon MM Host interfaces by 4.

```c
#include "HLS/hls.h"

compoment void
static_coalescing(ihc::mm_host<int, ihc::dwidth<128>, ihc::awidth<32>,
                   ihc::aspace<1>, ihc::latency<0>> &in,
                   ihc::mm_host<int, ihc::dwidth<128>, ihc::awidth<32>,
                   ihc::aspace<2>, ihc::latency<0>> &out,
                   int i) {

  // Four loads statically coalesced into one 16 bytes wide load
  int a1 = in[3 * i + 0];
  int a2 = in[3 * i + 1];
  int a3 = in[3 * i + 2];
  int a4 = in[3 * i + 3];

  // Four stores statically coalesced into one 16 bytes wide store
  out[3 * i + 0] = a4;
  out[3 * i + 1] = a3;
  out[3 * i + 2] = a2;
  out[3 * i + 3] = a1;
}
```

The System Viewer in the High-Level Design Reports for this example show that the design only has one load and one store, each of width 128 bit.
The compiler can infer a statically-coalesced aligned LSU if your coding style guarantees that your memory accesses will not be unaligned. Consider the following code example:

```c
component int alignedRead(AHost &A, unsigned int i)
{
    hls_register int memory_line[PARALLEL_WORDS] = {};
    int word_base = PARALLEL_WORDS * (i / PARALLEL_WORDS);
#pragma unroll
    for (int j = 0; j < PARALLEL_WORDS; j++)
    {
        int idx_base = PARALLEL_WORDS * ((i + j) / PARALLEL_WORDS);
        bool doesntWrap = idx_base == word_base;
        if (doesntWrap)
        {
            memory_line[j] = A[i + j];
        }
    }
...
}
The compiler infers a nonaligned LSU because the compiler cannot determine that all values of \( j \) are read from the same word of memory \( A \). Although the range of \( j \) matches exactly one word of memory, the accesses might wrap to another word. More than 1 read might be required. If you code your access with a base that aligns with the word width, the compiler can guarantee no wrapping occurs and so the compiler gives you an aligned LSU.

4.5. Agent Interfaces

The Intel HLS Compiler can implement two different types of agent interface for your component parameters: a control-and-status register (CSR) agent interface and a agent memory interface. In general, use the CSR agent interface to pass scalar values to your component and use the agent memory interface to pass large arrays to and from your component.

Agent interfaces are implemented as Avalon Memory Mapped (Avalon MM) Agent interfaces. For details about the Avalon MM Agent interfaces, see "Avalon Memory-Mapped Interfaces in Avalon Interface Specifications".

<table>
<thead>
<tr>
<th>Agent Type</th>
<th>Associated Agent Interface</th>
<th>Read/Write Behavior</th>
<th>Synchronization</th>
<th>Read Latency</th>
<th>Controlling Interface Data Width</th>
</tr>
</thead>
<tbody>
<tr>
<td>Register</td>
<td>The component CSR agent.</td>
<td>The component cannot update these registers from the datapath, so you can read back only data that you wrote in.</td>
<td>Synchronized with the component start signal.</td>
<td>Fixed value of 1.</td>
<td>Always 64 bits</td>
</tr>
<tr>
<td>Memory (M20K/MLAB)</td>
<td>Dedicated agent interface on the component.</td>
<td>The component reads from this memory and updates it as it runs. Updates from the component datapath are visible in memory.</td>
<td>Reads and writes to agent memories from outside of the component should occur only when your component is not executing. You might experience undefined component behavior if outside agent memory accesses occur when your component is executing. The undefined behavior can occur even if a agent memory access is to a memory address that the component does not access.</td>
<td>Fixed value that is dependent on the component memory access pattern and any attributes or pragmas that you set. See the Function Viewer report in the High-Level Design Report (report.html) for the read latency of a specific agent memory argument.</td>
<td>The data width is a multiple of the agent data type, where the multiple is determined by coalescing the internal accesses.</td>
</tr>
</tbody>
</table>

4.5.1. Control and Status Register (CSR) Agent

A component can have a maximum of one CSR agent interface, but more than one parameter can be mapped into this interface.

Any parameters that are labeled as hls_avalon_agent_register_argument are located in this memory space. The resulting memory map is described in the automatically generated header file `<results>.prj/components/`
<component_name>_csr.h. This file also provides the C macros for a host component to interact with the agent component. Examples of host components include Nios® II soft processors and Intel Acceleration Stack host applications.

The control and status registers (that is, function call and return) of a component with the hls_avalon_agent_component attribute are implemented in the CSR agent interface.

You do not need to use the hls_avalon_agent_component attribute to use the hls_avalon_agent_register_argument attribute.

To learn more, review the tutorial: <quartus_installdir>/hls/examples/tutorials/interfaces/mm_agents

Example code of a component with a CSR agent:

```c
#include "HLS/hls.h"

struct MyStruct {
    int f;
    double j;
    short k;
};

hls_avalon_agent_component

component MyStruct mycomp_xyz (hls_avalon_agent_register_argument int y,
    hls_avalon_agent_register_argument MyStruct struct_argument,
    hls_avalon_agent_register_argument unsigned long long mylong,
    hls_avalon_agent_register_argument char char_arg
) {
    return struct_argument;
}
```

Generated C header file for the component mycomp_xyz:

```c
/* This header file describes the CSR Agent for the mycomp_xyz component */

#ifndef __MYCOMP_XYZ_CSR_REGS_H__
#define __MYCOMP_XYZ_CSR_REGS_H__

/******************************************************************************/
/* Memory Map Summary                                                         */
/******************************************************************************/

/* Register  | Access  |   Register Contents      | Description */
Address   |         |      (64-bits)           |-----------------------------
------------|---------|--------------------------|-----------------------------
0x0 | R | {reserved[62:0], busy[0:0]} | Read the busy status of the component
0x8 | W | {reserved[62:0], start[0:0]} | Write 1 to signal start to the component
0x10 | R/W | {reserved[62:0], interrupt_enable[0:0]} | 0 - Disable interrupt, 1 - Enable interrupt
0x18 | R/Wclr | {reserved[61:0], done[0:0], interrupt_status[0:0]} | Signals component completion done is read-only and interrupt_status is write 1

4. Component Interfaces

MNL-1083 | 2021.06.21
44
## 4. Component Interfaces

### Register Address Macros

<table>
<thead>
<tr>
<th>Address</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0x0</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x8</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x10</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x20</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x28</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x30</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x38</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x40</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x48</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x50</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x58</td>
<td>Register Address Macros</td>
</tr>
<tr>
<td>0x60</td>
<td>Register Address Macros</td>
</tr>
</tbody>
</table>

NOTE: Writes to reserved bits will be ignored and reads from reserved bits will return undefined values.

/* Byte Addresses */

```c
#define MYCOMP_XYZ_CSR_BUSY_REG (0x0)
#define MYCOMP_XYZ_CSR_START_REG (0x8)
#define MYCOMP_XYZ_CSR_INTERRUPT_ENABLE_REG (0x10)
#define MYCOMP_XYZ_CSR_INTERRUPT_STATUS_REG (0x18)
#define MYCOMP_XYZ_CSR_RETURNDATA_0_REG (0x20)
#define MYCOMP_XYZ_CSR_RETURNDATA_1_REG (0x28)
#define MYCOMP_XYZ_CSR_RETURNDATA_2_REG (0x30)
#define MYCOMP_XYZ_CSR_ARG_Y_REG (0x38)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_0_REG (0x40)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_1_REG (0x48)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_2_REG (0x50)
#define MYCOMP_XYZ_CSR_ARG_MYLONG_REG (0x58)
#define MYCOMP_XYZ_CSR_ARG_CHAR_ARG_REG (0x60)
```

/* Argument Sizes (bytes) */

```c
#define MYCOMP_XYZ_CSR_RETURNDATA_0_SIZE (8)
#define MYCOMP_XYZ_CSR_RETURNDATA_1_SIZE (8)
#define MYCOMP_XYZ_CSR_RETURNDATA_2_SIZE (8)
#define MYCOMP_XYZ_CSR_ARG_Y_SIZE (4)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_0_SIZE (8)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_1_SIZE (8)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_2_SIZE (8)
#define MYCOMP_XYZ_CSR_ARG_MYLONG_SIZE (8)
#define MYCOMP_XYZ_CSR_ARG_CHAR_ARG_SIZE (1)
```

/* Argument Masks */

```c
#define MYCOMP_XYZ_CSR_RETURNDATA_0_MASK (0xffffffffffffffffULL)
#define MYCOMP_XYZ_CSR_RETURNDATA_1_MASK (0xffffffffffffffffULL)
#define MYCOMP_XYZ_CSR_RETURNDATA_2_MASK (0xffffffffffffffffULL)
#define MYCOMP_XYZ_CSR_ARG_Y_MASK (0xffffffff)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_0_MASK (0xffffffffffffffffull)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_1_MASK (0xffffffffffffffffull)
#define MYCOMP_XYZ_CSR_ARG_STRUCT_ARGUMENT_2_MASK (0xffffffffffffffffull)
#define MYCOMP_XYZ_CSR_ARG_MYLONG_MASK (0xffffffffffffffffull)
#define MYCOMP_XYZ_CSR_ARG_CHAR_ARG_MASK (0xff)
```
/* Status/Control Masks */
#define MYCOMP_XYZ_CSR_BUSY_MASK   (1<<0)
#define MYCOMP_XYZ_CSR_BUSY_OFFSET (0)
#define MYCOMP_XYZ_CSR_START_MASK   (1<<0)
#define MYCOMP_XYZ_CSR_START_OFFSET (0)
#define MYCOMP_XYZ_CSR_INTERRUPT_ENABLE_MASK   (1<<0)
#define MYCOMP_XYZ_CSR_INTERRUPT_ENABLE_OFFSET (0)
#define MYCOMP_XYZ_CSR_INTERRUPT_STATUS_MASK   (1<<0)
#define MYCOMP_XYZ_CSR_INTERRUPT_STATUS_OFFSET (0)
#define MYCOMP_XYZ_CSR_DONE_MASK   (1<<1)
#define MYCOMP_XYZ_CSR_DONE_OFFSET (1)
#endif /* __MYCOMP_XYZ_CSR_REGS_H__ */

4.5.2. Agent Memories

By default, component functions access parameters that are passed by reference through an Avalon Memory-Mapped (MM) Host interface. An alternative way to pass parameters by reference is to use an Avalon MM Agent memory interface, which exists inside the component.

Having a pointer argument generate an Avalon MM Host interface on the component has two potential disadvantages:

- The host interface has a single port. If the component has multiple load-store sites, arbitration on that port might create stallable logic.
- Depending on the system in which the component is instantiated, other hosts might use the memory bus while the component is running and create undesirable stalls on the bus.

Because a agent memory is internal to the component, the HLS compiler can create a memory architecture that is optimized for the access pattern of the component such as creating banked memories or coalescing memory accesses.

Agent memories differ from component memories because they can be accessed from an Avalon MM Host outside of the component. Component memories are by definition restricted to the component and cannot be accessed outside the component.

You can explicitly control the structure of your agent memories by applying memory arguments to agent memory variable declarations.

A component can have many agent memory interfaces. Unlike agent register arguments that are grouped together in the CSR agent interface, each agent memory has a separate interface with separate data buses. The agent memory interface data bus width is determined by the width of the agent type. If the internal accesses to the memory have been coalesced, the agent memory interface data bus width might be a multiple of the width of the agent type.

You can apply component memory attributes to agent memories in your component to customize the memory architecture and lower the FPGA area utilization of your component. For details, refer to Component Memory Attributes on page 55.
The following diagram shows a component with two agent memory interfaces defined as follows:

```c
component float myComponent(
    hls_avalon_agent_memory(64) float *a,
    hls_avalon_agent_memory(64) float *b) {
    return a[0]+b[0];
}
```

Reads and writes to agent memories from outside of the component should occur only when your component is not executing, unless you mark the agent memory argument with the `volatile` keyword. Without the `volatile` keyword, you might experience undefined component behavior if an external Avalon MM Host accesses your component agent memory when your HLS component is executing. The undefined behavior can occur even if the external access is to a memory address that your HLS component does not access.

**Volatile Agent Memories**

Add the `volatile` keyword to a agent memory argument to allow an Avalon MM Host to access the memory while the component executes. The compiler builds a memory system that ensures functional correctness even if the memory is modified from outside the component while the component is running.

However, allowing concurrent memory accesses with the `volatile` keyword might prevent some optimizations that the compiler can perform on your component. Consider this trade-off carefully in your design.
Depending on how the agent memory is accessed from within the component, the compiler might infer arbitration logic between the external memory access and internal memory accesses. In this case, the external access always gets arbitration priority over internal accesses. This prioritization might prevent the data flow in the component from progressing when external requests arrive at every cycle.

**Important:** You cannot verify the behavior of concurrent component memory access during simulation. To test the behavior, you must build a Verilog testbench to interact with your component. You can see an example of a Verilog testbench and how to use it in the following tutorial:

```<quartus_installdir>/hls/examples/tutorials/interfaces/mm_agents_CSR_volatile```

**Avalon MM Host Access Control**

You can indicate how an external Avalon MM Host interface accesses the component agent memory interface with the `hls_readwrite_mode` component macro:

- Use `hls_readwrite_mode("readonly")` to indicate that the external Avalon MM Host interface only ever reads from the agent memory. If you specify this macro, no write ports from outside of the component are created.
- Use `hls_readwrite_mode("writeonly")` to indicate that the external Avalon MM Host interface only ever writes to the agent memory. If you specify this macro, no read ports from outside of the component are created.

If you do not specify this macro, the compiler assumes that the external Avalon MM Host interface can read or write to the agent memory.

You can use the macro to help the compiler create a more efficient memory system and potentially save FPGA area.

You can use the `hls_readwrite_mode` macro for both volatile and non-volatile agent memories.

The following example shows how you can apply the macro:

```component void
foo(hls_avalon_agent_memory_argument(128*sizeof(int))
hls_readwrite_mode("writeonly") int *A)```

**Agent Memory Component Macros**

<table>
<thead>
<tr>
<th>Component Macro</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>hls_avalon_agent_memory_argument</code></td>
<td>Implement the parameter, in on-chip memory blocks, which can be read from or written to over a dedicated agent interface.</td>
</tr>
<tr>
<td><code>hls_readwrite_mode</code></td>
<td>Indicate to the compiler how the agent memory interface is accessed by external Avalon memory-mapped (MM) hosts.</td>
</tr>
</tbody>
</table>

**4.6. Stable Component Parameters**

If you do not specify the intended behavior for a parameter, the compiler assumes that the argument corresponding to a parameter changes value between pipelined component invocations.
If the corresponding argument for a parameter does not change while your component executes, you can mark the parameter as stable. The arguments for a stable parameter can still change after all active component invocations have finished (that is, the component datapath is flushed). If an argument changes while there are active component invocations in progress, your component enters an undefined state and behaves unpredictably.

Declare an interface parameter to be stable with the `hls_stable_argument` attribute.

You can mark the following the interface parameters as stable:

- Scalar (conduit) parameters
- Pointer interface parameters
  The address conduit input is stable. The associated Avalon MM Host interface is not affected.
- Pass-by-reference parameters
  The address conduit input is stable. The associated Avalon MM Host interface is not affected.
- Avalon Memory-Mapped (MM) Host interface parameters
  The address conduit input is stable. The associated Avalon MM Host interface is not affected.
- Avalon Memory-Mapped (MM) Agent register interface parameters
  The data is stored in the CSR, but not registered in the component data path.

The following interface parameters cannot be marked as stable:

- Avalon Memory-Mapped (MM) Agent memory interface parameters
- Avalon Streaming interface parameters

You might save some FPGA area in your component design when you declare an interface argument as stable because the data is not registered throughput.

You cannot have two component invocations in flight with different stable arguments between the two component invocations.

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>hls_stable_argument</code></td>
<td>A stable parameter is a parameter that does not change while there is live data in the component (that is, the argument does not change between pipelined function invocations).</td>
</tr>
</tbody>
</table>

**4.7. Global Variables**

Components can use and update C++ global variables. If you access a global variable in your component function, it is implemented as an Avalon Memory-Mapped (MM) Host interface, like a pointer parameter.

If you access more than one global variable, each global variable uses the same Avalon MM Host interface, which results in stallable arbitration. If you use pointers and non-constant global memory accesses, then the pointers and global memory accesses all share the same Avalon MM Host interface.
In addition to the Avalon MM Host interface, each global variable that the component uses has an input conduit that must be supplied with the address of the global variable in system memory. The input conduit arguments that are generated in the RTL are named @<global variable name>. Input conduits generated for pointer arguments omit the @ are named for the corresponding pointer argument.

If your global variable is declared as const, then no Avalon MM Host interface and no additional input conduit is generated. Therefore, global variables declared as const use significantly less FPGA area than modifiable global variable.

4.8. Structs in Component Interfaces

Review the interface_structs.sv file in your <a.prj>/components/<component_name> folder to see information about the padding and packed-ness of the implementation interfaces for the structs in your component.

The interface_structs.sv file contains the Verilog-style definitions of the structs found on your component interface.

4.9. Reset Behavior

For your HLS component, the reset assertion can be asynchronous but the reset deassertion must be synchronous.

The reset assertion and deassertion behavior can be generated from an asynchronous reset input by using a reset synchronizer, as described in the following example Verilog code:

```verilog
reg [2:0] sync_resetn;
always @(posedge clock or negedge resetn) begin
    if (!resetn) begin
        sync_resetn <= 3'b0;
    end else begin
        sync_resetn <= {sync_resetn[1:0], 1'b1};
    end
end
wire synchronized_resetn;
assign synchronized_resetn = sync_resetn[2];
```

This synchronizer code is used in the example Intel Quartus Prime Pro Edition project that is generated for your components included in an i++ compile.

When the reset is asserted, the component holds its busy signal high and its done signal low. After the reset is deasserted, the component holds its busy signal high until the component is ready to accept the next invocation. All component interfaces (agents, hosts, and streams) are ready only after the component busy signal is low.
**Simulation Component Reset**

You can check the reset behavior of your component during simulation by using the `ihc_hls_sim_reset` API. This API returns 1 if the reset was exercised (that is, if the reset is called during hardware simulation of the component). Otherwise, the API returns 0.

Call the API as follows:

```c
int ihc_hls_sim_reset(void);
```

During x86 emulation of your component, the `ihc_hls_sim_reset` API always returns 0. You cannot reset a component during x86 emulation.
5. Component Memories (Memory Attributes)

The Intel High Level Synthesis (HLS) Compiler can build a hardware memory system using FPGA memory resources (such as block RAMs) for local, constant, and static variables as well as agent memory declared in your code.

In some cases, the Intel HLS Compiler implements a local, constant, and static variable using registers in the component datapath. However, you can override that implementation by using memory attributes.

Memory accesses are mapped to load-store units (LSUs), which transact with the hardware memory through its ports. The Intel HLS Compiler sometimes statically coalesces multiple memory accesses to a component memory into one wider memory access in order to save on the number of LSUs instantiated. LSUs for component memory are always pipelined LSUs.

If two or more LSUs need to be scheduled during the same cycle, the compiler might create stallable arbitration logic. Stallable arbitration logic appears red in the Component Viewer (in the High-Level Design Reports).

For more details about LSUs instantiated by the Intel HLS Compiler, see Load-Store Unit Types on page 36. For details about coalescing memory accesses to save on instantiated LSUs, see Memory-Access Coalescing and Load-Store Units on page 40.

Figure 7. A Basic Memory Configuration Inferred by the Intel HLS Compiler

The following diagram shows a basic memory configuration:

```
Memory Size = 4096 bytes

Bank 0 (1024x4)
Replicate 0 (1024x4)
Copy 0 (1024x4)

Depth = 1024 words

Width = 4 bytes
```

Intel Corporation. All rights reserved. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Intel warrants performance of its FPGA and semiconductor products to current specifications in accordance with Intel's standard warranty, but reserves the right to make changes to any products and services at any time without notice. Intel assumes no responsibility or liability arising out of the application or use of any information, product, or service described herein except as expressly agreed to in writing by Intel. Intel customers are advised to obtain the latest version of device specifications before relying on any published information and before placing orders for products or services.

*Other names and brands may be claimed as the property of others.*
Figure 8. A Memory System With Two Memory Banks

The contents of a memory system can be partitioned into one or more memory banks, such that each bank contains a subset of data contained in the hardware memory:

- **Memory Size = 8192 bytes**
- **Depth = 2048 words**
- **Width = 4 bytes**
A memory bank can contain one or more memory replicates. The compiler might create memory replicates to create more read ports. Having more read ports allows concurrent access to your memory system if you have many read operations.

The replicates in a memory bank contain identical data and you can read from the replicates simultaneously. A replicate can have two or four access ports, depending on whether the replicate is clocked at the same frequency (single pumped) or twice the frequency (double pumped) of the component. All ports in replicates can be accessed concurrently. The number of ports in a memory bank depends on the number of replicates that the bank contains.

A replicate can also contain one or more private copies to support multiple concurrent loop iterations.
The Intel HLS Compiler can control the geometry and configuration parameters of the hardware memories that it builds. The compiler tries to create stall-free memory accesses. That is, the compiler tries to give memory reads and writes contention-free access to a memory port. A memory system is stall-free if all reads and writes in the memory system are contention-free.

The compiler tries to create a minimum-area stall-free memory system. If you want a different area-performance trade off, use the component memory attributes to specify your own memory system configuration and override the memory system inferred by the compiler.

**Component Memory Attributes**

Apply the component memory attributes to local, constant, and static variables in your component to customize the on-chip memory architecture of the component memory system and lower the FPGA area utilization of your component. You can also apply memory attributes to agent memories and struct data members.

These component memory attributes are defined in the "HLS/hls.h" header file, which you can include in your code.

<table>
<thead>
<tr>
<th>Memory Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hls_force_pow2_depth</td>
<td>Specifies that the memory implementing the variable or array has power-of-2 depth.</td>
</tr>
<tr>
<td>hls_register</td>
<td>Forces a variable or array to be carried through the pipeline in registers. A register variable can be implemented either exclusively in flip-flops (FFs) or in a mix of FFs and RAM-based FIFOs.</td>
</tr>
</tbody>
</table>

**Figure 10. A Memory System With Two Private Copies**

The memory system shown in Figure 10 has two private copies. The memory size is 8192, with Bank 0 (2048x4) and Replicate 0 (2048x4). The depth is 2048 words and the width is 4 bytes.
Memory Attribute | Description
--- | ---
hls_memory | Forces a variable or array to be implemented as embedded memory.
hls_memory_impl | Forces a variable or array to be implemented as embedded memory of a specified type.
hls_singlepump | Specifies that the memory implementing the variable or array must be clocked at the same rate as the component accessing the memory.
hls_doublepump | Specifies that the memory implementing the variable or array must be clocked at twice the rate as the component accessing the memory.
hls_numbanks | Specifies that the memory implementing the variable or array must have a defined number of memory banks.
hls_bankwidth | Specifies that the memory implementing the variable or array must have memory banks of a defined width.
hls_bankbits | Forces the memory system to split into a defined number of memory banks and defines the bits used to select a memory bank.
hls_simple_dual_port_memory | Specifies that the memory implementing the variable or array should have no port that services both reads and writes.
hls_merge (depthwise) | Allows merging two or more local variables to be implemented in component memory as a single merged memory system in a depth-wise manner.
hls_merge (widthwise) | Allows merging two or more local variables to be implemented in component memory as a single merged memory system in a width-wise manner.
hls_init_on_reset | Forces the static variables inside the component to be initialized when the component reset signal is asserted.
hls_init_on_powerup | Sets the component memory implementing the static variable to initialize on power-up when the FPGA is programmed.
hls_max_concurrency | Deprecated: This attribute is deprecated and will be removed in a future release. Use the hls_private_copies memory attribute instead. Specifies that the memory implementing the variable or array has a defined number of private copies to allow concurrent iterations of a loop at any given time.
hls_max_replicates | Specifies that the memory implementing the variable or array has no more than the specified number of replicates to enable simultaneous reads from the datapath.
hls_private_copies | Specifies that the memory implementing the variable or array has a defined number of private copies to allow concurrent iterations of a loop at any given time.

**Struct Datatypes and Memory Attributes**

You can apply memory attributes to `struct` member variables in the `struct` declaration. If you also apply memory attributes to the object instantiation of a `struct` variable, the attributes on the instantiation override the attributes from the declaration.

The following code example applies memory attributes to both a declaration and an instantiation:

```c
struct State {
    int array[100] hls_memory;
    int reg[4] hls_register;
};
```
component int test(..) {
    struct State S1;
    struct State S2 hls_memory;
    // some uses
}

For this example code, the compiler splits \texttt{S1} into two variables, \texttt{S1.array[100]} (implemented in memory) and \texttt{S1.reg[4]} (implemented in registers). However, the compiler ignores the attributes applied at the \texttt{struct} declaration for object \texttt{S2} because the \texttt{S2} object has the \texttt{hls_memory} attribute applied at instantiation.

**Constraints on Attributes for Memory Banks**

The properties of memory banks constrain how you can divide component memory into banks with the memory bank attributes.

The relationship between the following properties is constrained:

- The number of bytes in your array that you want to access at one time (\(S\)). If you are accessing a local variable, this value represents the size (in bytes) of the local variable.
- The number of memory banks specified by \texttt{hls_numbanks} attribute (\(N_{\text{banks}}\)).
- The width (in bytes) of the memory banks specified by \texttt{hls_bankwidth} attribute (\(W\)).
- The number of memory bank-select bits specified by \texttt{hls_bankbits} attribute. That is, \(n+1\) when you specify \(b_0, b_1, ..., b_n\) as the bank-select bits (\(N_{\text{bits}}\)).

These attributes are subject to the following constraints:

- \(N_{\text{banks}} \times W = S\)
  
The number of bytes accessed concurrently (or size of a local variable) is equal to the number of memory banks it uses times the width of the memory banks.
- \(N_{\text{banks}}\) must be a power of 2 value.
- \(N_{\text{banks}} = 2^{N_{\text{bits}}}\)

\(N_{\text{bits}}\) bank-selection bits that are required to address \(N_{\text{banks}}\) number of memory banks.

Values that you specify for the \texttt{hls_numbanks}, \texttt{hls_bankwidth}, and \texttt{hls_bankbits} attributes must meet these constraints. For attributes that you do not specify, the Intel HLS Compiler infers values for the attributes following these constraints.

To learn more about these attributes, review the following tutorial:

\(<\text{quartus_installdir}>/\text{hls/examples/tutorials/component_memories/memory\_geometry}\>

**Related Information**

Agent Memories on page 46
5.1. Static Variables

The HLS compiler supports function-scope static variables with the same semantics as in C and C++. Function-scope static variables are initialized to the specified values on reset. In addition, changes to these variables are visible across component invocations, making function-scope static variables ideal for storing state in a component. However, function-scope static variables cannot be shared by different task or component functions.

To initialize static variables, the component requires extra logic, and the component might take some time to exit the reset state while this logic is active.

**Static Variable Initialization**

You can control when the static variables in your component are initialized if the variables are implemented as memories. The memory system that stores a static variable can be initialized either when your component is powered up or when your component is reset.

Initializing a static variable when a component is powered up resembles a traditional programming model where you cannot reinitialize the static variable value after the program starts to run.

Initializing a static variable when a component is reset initializes the static variable each time your component receives a reset signal, including on power up. However, this type of static variable initialization requires extra logic. This extra logic can affect the start-up latency and the FPGA area needed for your component.

You can explicitly set the static variable initialization by adding one of the following attributes to your static variable declaration:

- **hls_init_on_reset**
  - The static variable value is initialized after the component is reset.
  - Add this attribute to your static variable declaration as shown in the following example:
    ```
    static char arr[128] hls_init_on_reset;
    ```
  - This is the default behavior for initializing static variables. You do not need to specify the `hls_init_on_reset` keyword with your static variable declaration to get this behavior.
  - For example, the static variable in the following example is initialized when the component is reset:
    ```
    static int arr[64];
    ```

- **hls_init_on_powerup**
  - The static variable is initialized only on power up. This initialization uses a memory initialization file (.mif) to initialize the memory, which reduces the resource utilization and start-up latency of the component.
Add this keyword to your static variable declaration as shown in the following example:

```c
static char arr[128] hls_init_on_powerup;
```

Some static variables might not be able to take advantage of this initialization because of the complexity of the static variables (for example, an array of structs). In these cases, the compiler returns an error.

For a demonstration of initializing static variables, review the tutorial in `<quartus_installdir>/hls/examples/tutorials/component_memories/static_var_init`.

For information about resetting your component, see Reset Behavior on page 50.
6. Loops in Components

The Intel HLS Compiler Pro Edition attempts to pipeline loops to maximize throughput of the various components that you define.

Loop Pipelining

Pipelining loops enables the Intel HLS Compiler Pro Edition to execute subsequent iterations of a loop in a pipeline-parallel fashion. Pipeline-parallel execution means that multiple iterations of the loop, at different points in their executions, are executing at the same time. Because all stages of the loop are always active, pipelining loops helps maximize usage of the generated hardware. This loop pipelining is similar to the hardware pipelining described in Intel HLS Compiler Hardware Model on page 14.

Figure 11. Pipelined loop with three stages and four iterations

In this figure, one stage is the logic that runs during one clock cycle.

There are some cases where pipelining is not possible at all. In other cases, a new iteration of the loop cannot start until N cycles after the previous iteration.

The number of cycles for which a loop iteration must wait before it can start is called the initiation interval (II) of the loop. This loop pipelining status is captured in the high level design report (report.html). In general, an II of 1 is desirable.

A common case where II > 1 is when a part of the loop depends in some way on the results of the previous iteration of the same loop. The circuit must wait for these loop-carried dependencies to be resolved before starting a new iteration of the loop. These loop-carried dependencies are indicated in the optimization report.
In the case of nested loops, II > 1 for an outer loop is not considered a significant performance limiter if a critical inner loop carries out the majority of the work. One common performance limiter is if the HLS compiler cannot statically compute the trip count of an inner loop (for example, a variable inner loop trip count). Without a known trip count, the compiler cannot pipeline the outer loop.

For more information about loop pipelining, see Pipeline Loops in Intel High Level Synthesis Compiler Best Practices Guide.

**Compiler Pragmas Controlling Loop Pipelining**

The Intel HLS Compiler has several pragmas that you can specify in your code to control how the compiler pipelines your loops.

Most loop pragmas must immediately precede the loop that the pragma applies to. You cannot have a loop pragma before elements such as labels on loops. The following table shows examples of how to apply loop pragmas correctly.

<table>
<thead>
<tr>
<th>Incorrect (produces a compile-time error)</th>
<th>Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>#pragma ivdep</td>
<td>TEST_LOOP: for(int idx = 0; idx &lt; counter; idx++) {...}</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pragma</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>disable_loop_pipelining</td>
<td>Prevents compiler from pipelining a loop,</td>
</tr>
<tr>
<td>ii</td>
<td>Forces a loop to have a loop initiation interval (II) of a specified value.</td>
</tr>
<tr>
<td>ivdep</td>
<td>Ignores memory dependencies between iterations of this loop.</td>
</tr>
<tr>
<td>loop_coalesce</td>
<td>Tries to fuse all loops nested within this loop into a single loop.</td>
</tr>
<tr>
<td>loop_fuse</td>
<td>Directs the compiler to try and fuse pairs of adjacent loops.</td>
</tr>
<tr>
<td>max_concurrency</td>
<td>Limits the number of iterations of a loop that can simultaneously execute at any time.</td>
</tr>
<tr>
<td>max_interleaving</td>
<td>Controls whether iterations of a pipelined inner loop in a loop nest from one invocation of the inner loop can be interleaved in the component data pipeline with iterations from other invocations of the inner loop.</td>
</tr>
<tr>
<td>nofusion</td>
<td>Prevents the annotated loop from being fused with adjacent loops.</td>
</tr>
<tr>
<td>speculated_iterations</td>
<td>Specifies the number of clock cycles that a loop exit condition can take to compute.</td>
</tr>
<tr>
<td>unroll</td>
<td>Unrolls the loop completely or by a number of times.</td>
</tr>
</tbody>
</table>

For a list of tutorials that demonstrate best practices to follow when implementing loops and using the loop pragmas in your component, see Loop Best Practices in Intel High Level Synthesis Compiler Best Practices Guide.
6.1. Loop Initiation Interval (\texttt{ii} Pragma)

The initiation interval, or II, is the number of clock cycles between the launch of successive loop iterations. Use the \texttt{ii} pragma to direct the Intel High Level Synthesis (HLS) Compiler to attempt to set the initiation interval (II) for the loop that follows the pragma declaration. If the compiler cannot achieve the specified II for the loop, then the compilation errors out.

You might want to increase the II of a loop to get an \texttt{f}_{\text{MAX}} improvement in your component. A loop is a good candidate to have the \texttt{ii} pragma applied to increase its loop II if the loop meets any of the following conditions:

- The loop is not critical to the throughput of your component.
- The running time of the loop is small compared to other loops it might contain.

You can also apply the \texttt{ii} pragma to force a loop to an II of 1 and accept a possible \texttt{f}_{\text{MAX}} penalty.

To specify a loop initiation interval for a loop, specify the pragma before the loop as follows:

```c
#pragma ii <desired_initiation_interval>
```

The \texttt{<desired_initiation_interval>} parameter is required and is an integer that specifies the number of clock cycles to wait between the beginning of execution of successive loop iterations.

You can also direct the compiler to attempt to achieve an II for your component with the \texttt{hls\_component\_ii}\ (<\texttt{N}>\) component attribute. For more information, refer to \texttt{hls\_component\_ii}.

To learn more about the effects of applying the \texttt{ii} loop pragma, review the following tutorial:

```c
<quartus\_install\_dir>/hls/examples/tutorials/best\_practices/set\_target\_fmax_1
```

6.2. Loop-Carried Dependencies (\texttt{ivdep} Pragma)

When compiling your components, the HLS compiler generates hardware to avoid any data hazards between load and store instructions to component memories, agent memories, and external memories (through Avalon-MM host interfaces). In particular, read-write dependencies can limit performance when they exist across loop iterations because they prevent the compiler from beginning a new loop iteration before the current iteration finishes executing its load and store instructions. You have the option to guarantee to the HLS compiler that there are no implicit memory dependencies across loop iterations in your component by adding the \texttt{ivdep} pragma in your code.

The \texttt{ivdep} pragma tells the compiler that a memory dependency between loop iterations can be ignored. Ignoring the dependency saves area and lowers the loop initiation interval (II) of the affected loop because the hardware required for avoiding data hazards is no longer required.

You can provide more information about loop dependencies by adding the \texttt{safelen(N)} clause to the \texttt{ivdep} pragma. The \texttt{safelen(N)} clause specifies the maximum number of consecutive loop iterations without loop-carried memory dependencies. For example, \#pragma \texttt{ivdep safelen(32)} indicates to the compiler...
that there are a maximum of 32 iterations of the loop before loop-carried dependencies might be introduced. That is, while #pragma ivdep promises that there are no implicit memory dependency between any iteration of this loop, #pragma safelen(32) promises that the iteration that is 32 iterations away is the closest iteration that could depend on this iteration.

To specify that accesses to a particular memory array inside a loop will not cause loop-carried dependencies, add the line #pragma ivdep array (array_name) before the loop in your component code. The array specified by the ivdep pragma must be one of the following items:

- a component memory array
- a pointer argument
- a pointer variable that points to a component memory
- a reference to an mm_host object

If the specified array is a pointer, the ivdep pragma also applies to all arrays that may alias with specified pointer. The array specified by the ivdep pragma can also be an array or a pointer member of a struct.

**Caution:** Incorrect usage of the ivdep pragma might introduce functional errors in hardware.

### Use Case 1:
If all accesses to memory arrays inside a loop do not cause loop-carried dependencies, add #pragma ivdep before the loop.

```c
1  // no loop-carried dependencies for A and B array accesses
2  #pragma ivdep
3  for(int i = 0; i < N; i++) {
4      A[i] = A[i + N];
5      B[i] = B[i + N];
6  }
```

### Use Case 2:
You may specify #pragma ivdep array (array_name) on particular memory arrays instead of all array accesses. This pragma is applicable to arrays, pointers, or pointer members of structs. If the specified array is a pointer, the ivdep pragma applies to all arrays that may alias with the specified pointer.

```c
1  // No loop-carried dependencies for A array accesses
2  // Compiler inserts hardware that reinforces dependency constraints for B
3  #pragma ivdep array(A)
4  for(int i = 0; i < N; i++) {
5      A[i] = A[i - X[i]];  
6      B[i] = B[i - Y[i]];  
7  }
8
9  // No loop-carried dependencies for array A inside struct
10  #pragma ivdep array(S.A)
11  for(int i = 0; i < N; i++) {
13  }
14
15  // No loop-carried dependencies for array A inside the struct pointed by S
16  #pragma ivdep array(S->X[2][3].A)
17  for(int i = 0; i < N; i++) {
19  }
```
// No loop-carried dependencies for A and B because ptr aliases with both arrays
int *ptr = select ? A : B;

#pragma ivdep array(ptr)
for(int i = 0; i < N; i++) {
    A[i] = A[i - X[i]];
    B[i] = B[i - Y[i]];
}

// No loop-carried dependencies for A because ptr only aliases with A
int *ptr = &A[10];

#pragma ivdep array(ptr)
for(int i = 0; i < N; i++) {
    A[i] = A[i - X[i]];
    B[i] = B[i - Y[i]];
}

To learn more about the effects of applying the ivdep loop pragma, review the following tutorial:

<quartus_installdir>/hls/examples/tutorials/best_practices/loop_memory_dependency

6.3. Loop Coalescing (loop_coalesce Pragma)

Use the loop_coalesce pragma to direct the Intel HLS Compiler to coalesce nested loops into a single loop without affecting the loop functionality. Coalescing loops can help reduce your component area usage by directing the compiler to reduce the overhead needed for loop control.

Coalescing nested loops also reduces the latency of the components, which can further reduce your component area usage. However, in some cases, coalescing loops might lengthen the critical loop initiation interval path, so coalescing loops might not be suitable for all components.

To coalesce nested loops, specify the pragma as follows:

#pragma loop_coalesce <loop_nesting_level>

The <loop_nesting_level> parameter is optional and is an integer that specifies how many nested loop levels that you want the compiler to attempt to coalesce. If you do not specify the <loop_nesting_level> parameter, the compiler attempts to coalesce all of the nested loops.

For example, consider the following set of nested loops:

for (A)
    for (B)
        for (C)
            for (D)
                for (E)

If you place the pragma before loop (A), then the loop nesting level for these loops is defined as:
• Loop (A) has a loop nesting level of 1.
• Loop (B) has a loop nesting level of 2.
• Loop (C) has a loop nesting level of 3.
• Loop (D) has a loop nesting level of 4.
• Loop (E) has a loop nesting level of 3.

Depending on the loop nesting level that you specify, the compiler attempts to coalesce loops differently:

• If you specify #pragma loop_coalesce 1 on loop (A), the compiler does not attempt to coalesce any of the nested loops.
• If you specify #pragma loop_coalesce 2 on loop (A), the compiler attempts to coalesce loops (A) and (B).
• If you specify #pragma loop_coalesce 3 on loop (A), the compiler attempts to coalesce loops (A), (B), (C), and (E).
• If you specify #pragma loop_coalesce 4 on loop (A), the compiler attempts to coalesce all of the loops [loop (A) - loop (E)].

To learn more about the effects of applying the `loop_coalesce` loop pragma, review the following tutorial:

<quartus_installdir>/hls/examples/tutorials/best_practices/loop_coalesce

**Example**

The following simple example shows how the compiler coalesces two loops into a single loop.

Consider a simple nested loop written as follows:

```c
#pragma loop_coalesce
for (int i = 0; i < N; i++)
    for (int j = 0; j < M; j++)
        sum[i][j] += i+j;
```

The compiler coalesces the two loops together so that they run as if they were a single loop written as follows:

```c
int i = 0;
int j = 0;
while(i < N){
    sum[i][j] += i+j;
    j++;
    if (j == M){
        j = 0;
        i++;
    }
}
```

### 6.4. Loop Unrolling (unroll Pragma)

The Intel HLS Compiler supports the `unroll` pragma for unrolling multiple copies of a loop.
Example code:

```c
#pragma unroll <N>
for (int i = 0; i < M; ++i) {
    // Some useful work
}
```

In this example, `<N>` specifies the unroll factor, that is, the number of copies of the loop that the HLS compiler generates. If you do not specify an unroll factor, the HLS compiler unrolls the loop fully when the number of loop iterations is known at compile time.

As an example of the kind of code the compiler generates when unrolling a loop, consider the following code snippet where you specify an unroll factor of 3:

```c
hls_register float data[N];
#pragma unroll 3
for (int i = 0; i < N; i++)
{
    data[i] = function(i, a);
}
```

Unrolling the loop with an unroll factor of 3 results in the compiler transforming the code snippet into something like the following code:

```c
hls_register float data[N];
for (int i = 0; i < N; i += 3)
{
    data[i + 0] = function(i + 0, a);
    if (i + 1 < N)
    {
        data[i + 1] = function(i + 1, a);
    }
    if (i + 2 < N)
    {
        data[i + 2] = function(i + 2, a);
    }
}
```

You can find the unroll status of each loop in the high level design report (report.html).

### 6.5. Loop Concurrency (max_concurrency Pragma)

You can use the `max_concurrency` pragma to increase or limit the concurrency of a loop in your component. The concurrency of a loop is how many iterations of that loop can be in progress at one time. By default, the Intel HLS Compiler tries to maximize the concurrency of loops so that your component runs at peak throughput.

To achieve maximum concurrency in loops, sometimes private copies of component memory have to be created to break dependencies on the underlying hardware that prevent the loop from being fully pipelined.

You can see the number of private copies created for you component memories in the High-Level Design reports (report.html) for your component:
In the Details pane of the Loop Analysis report as a message that says that the maximum number of simultaneous executions has been limited to N.

In the Bank view of your component memory in the Function Memory Viewer, which graphically shows the number of private copies.

Creating private copies of component memory in this case is not the same as replicating memory in order to increase the number of ports.

If you want to exchange some performance for component memory savings, apply #pragma max_concurrency <N> to the loop. When you apply this pragma, the number of private copies changes and controls the number of iterations entering the loop, as shown in the following example:

```c
#pragma max_concurrency 1
for (int i = 0; i < N; i++) {
    int arr[M];
    // Doing work on arr
}
```

You can control the number of private copies created for a component memory accessed withing a loop by using the hls_private_copies memory attribute. For details, see hls_private_copies Memory Attribute. For an example of increasing the number of private copies of memory, refer to the following tutorial:

```bash
<quartus_installdir>/hls/examples/tutorials/component_memories/non_power_of_two_memory
```

You can also control the concurrency of your component by using the hls_max_concurrency component attribute. For more information about the hls_max_concurrency(N) component attribute, see Concurrency Control (hls_max_concurrency Attribute). For an example increasing concurrency for better throughput, refer to the following tutorial:

```bash
<quartus_installdir>/hls/examples/tutorials/best_practices/optimize_ii_using_hls_register
```

### 6.6. Loop Iteration Speculation (speculated_iterations Pragma)

With speculated_iterations pragma control, you can adjust the number of speculated iterations for a loop. Speculated iterations are loop iterations that are initiated while the loop exit condition is being calculated. Adjusting the number of speculated iterations can help enable more efficient loop pipelining in your component.

The Intel HLS Compiler determines the number of speculated iterations on a per-loop basis. You can see the number of speculated iterations for a loop in the Loop Analysis Report in the High Level Design Report (report.html).

For loops where the exit condition calculation is a bottleneck (as shown in the Loop Analysis Report), consider increasing the number of speculated iterations with the speculated_iterations pragma. Increasing the number of speculated iterations might not improve the loop II if other bottlenecks in the loop are found.
For frequently invoked loops with a low latency loop body (for example, an inner loop with a short trip count), you might want to use the `speculated_iterations` pragma to reduce the number of speculated iterations to reduce the overhead of your design. However, setting the number of speculated iterations too low might increase the loop II because there is not enough time to evaluate the exit condition.

For examples of the effects of using the `speculated_iterations` pragma, review the following tutorials:

```
<quartus_installdir>/hls/examples/tutorials/loop_controls/speculated_iterations
<quartus_installdir>/hls/examples/tutorials/loop_controls\small_speculated_iterations
```

**Related Information**

"Speculative Execution" in the *Intel HLS Compiler Pro Edition Best Practices Guide*

### 6.7. Loop Pipelining Control (disable_loop_pipeliningPragma)

When the loop iterations effectively execute sequentially due to loop-carried dependencies, use the `disable_loop_pipelining` pragma to generate a simple sequential datapath and avoid loop resource hardware duplication. The simpler datapath and lack of resource duplication in hardware reduces the FPGA area utilization of your component.

Use the **Loop Analysis** section of the high-level design reports (*report.html*) to help determine if you should apply this pragma to your loops.

In the following example, the Intel HLS Compiler fails to schedule the loop with a small loop initiation interval (II) because of a memory dependency. Pipelining this loop is unlikely to have any benefit to your component throughput or performance.

```
#pragma disable_loop_pipelining
for (int i = 1; i < N; i++) {
    int j = a[i-1];
    // Memory dependency induces a high-latency loop feedback path
    a[i] = foo(j)
}
```

You can also disable pipelining the datapath of your entire component with the `hls_disable_component_pipelining` component attribute. For more information about this attribute, see **Component Pipelining Control (hls_disable_component_pipelining Attribute)** on page 78.

### 6.8. Loop Interleaving Control (max_interleavingPragma)

The Intel HLS Compiler Pro Edition tries to maximize the throughput and hardware resource occupancy of pipelined inner loops in a loop nest by issuing new inner loop iterations as frequently as possible (minimizing the loop initiation interval). When the compiler cannot achieve a loop II of 1 for an inner loop, the compiler configures the loop nest to interleave iterations of one invocation of the inner loop with iterations of other invocations of the inner loop.

**Terminology**

**Reminder**

A loop iteration is the single execution of a loop body. A loop invocation is the start of pipelined execution of loop iterations.
Interleaving Example 1

As another example, consider the following loop nest:

```c
// Loop j is pipelined with ii=1
for (int j = 0; j < M; j++) {
  int a[N];
  // Loop i is pipelined with ii=2
  for (int i = 1; i < N; i++) {
    a[i] = foo(i)
  }
}
```

In this example, the inner loop \( i \) is pipelined with \( \text{II}=2 \). Under normal pipelining, this II means that the inner loop hardware only achieves 50% utilization, since one iteration of the \( i \) loop is initiated every other cycle. To take advantage of these idle cycles, the compiler interleaves a second invocation of the \( i \) loop from the next iteration of the outer \( j \) loop.

Because the \( i \) loop resides inside the \( j \) loop, and the \( j \) loop has a trip count of \( M \), the \( i \) loop is *invoked* \( M \) times. The \( j \) loop is the outermost loop and is invoked once.

The following table shows the difference between normal pipelined execution of the \( i \) loop versus interleaved execution for this example for \( N=5 \).

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Pipelined Loop Iterations (j loop, i loop)</th>
<th>Interleaved Loop Iterations (j loop, i loop)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>(0,0)</td>
<td>(0,0)</td>
</tr>
<tr>
<td>1</td>
<td>---</td>
<td>(1,0)</td>
</tr>
<tr>
<td>2</td>
<td>(0,1)</td>
<td>(0,1)</td>
</tr>
<tr>
<td>3</td>
<td>---</td>
<td>(1,1)</td>
</tr>
</tbody>
</table>

---

Interleaving Example 2
This table shows the values \((j,i)\) for each inner loop iteration that is initiated at each cycle. At cycle 0, both modes of execution initiate the \((0,0)\)th iteration of the \(i\) loop. Under normal pipelined execution, no \(i\) loop iteration is initiated at cycle 1. Under interleaved execution, the \((1,0)\)th iteration of the innermost loop, i.e. the first iteration of the next \((j=1)\) invocation of the \(i\) loop, is initiated. By cycle 10, interleaved execution has initiated all of the iterations of both the \(j=0\) invocation of the \(i\) loop, and the \(j=1\) invocation of the \(i\) loop. This represents twice the efficiency of the normal pipelined execution.

Sometimes you might determine that this interleaving does not give you a performance benefit relative to the additional FPGA area needed to enable interleaving. In these cases, you can limit or restrict the amount of interleaving to reduce FPGA area utilization.

### Using the max_interleavingPragma

To limit the number of interleaved invocations of an inner loop that can be executed simultaneously, annotate the inner loop with the `max_interleaving` pragma. The annotated loop must be contained inside another pipelined loop.

The required parameter \((n)\) specifies an upper bound on the degree of interleaving allowed, That is, how many invocations of the containing loop can execute the annotated loop at a given time.

Specify the `max_interleaving` pragma in one of the following ways:

<table>
<thead>
<tr>
<th>Cycle</th>
<th>Pipelined Loop Iterations ((j \text{ loop, } i \text{ loop}))</th>
<th>Interleaved Loop Iterations ((j \text{ loop, } i \text{ loop}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>(0,2)</td>
<td>(0,2)</td>
</tr>
<tr>
<td>5</td>
<td>---</td>
<td>(1,2)</td>
</tr>
<tr>
<td>6</td>
<td>(0,3)</td>
<td>(0,3)</td>
</tr>
<tr>
<td>7</td>
<td>---</td>
<td>(1,3)</td>
</tr>
<tr>
<td>8</td>
<td>(0,4)</td>
<td>(0,4)</td>
</tr>
<tr>
<td>9</td>
<td>---</td>
<td>(1,4)</td>
</tr>
<tr>
<td>10</td>
<td>(1,0)</td>
<td>(2,0)</td>
</tr>
<tr>
<td>11</td>
<td>---</td>
<td>(3,0)</td>
</tr>
<tr>
<td>12</td>
<td>(1,1)</td>
<td>(2,1)</td>
</tr>
<tr>
<td>13</td>
<td>---</td>
<td>(3,1)</td>
</tr>
<tr>
<td>14</td>
<td>(1,2)</td>
<td>(2,2)</td>
</tr>
<tr>
<td>15</td>
<td>---</td>
<td>(3,2)</td>
</tr>
<tr>
<td>16</td>
<td>(1,3)</td>
<td>(2,3)</td>
</tr>
<tr>
<td>17</td>
<td>---</td>
<td>(3,3)</td>
</tr>
<tr>
<td>18</td>
<td>(1,4)</td>
<td>(2,4)</td>
</tr>
<tr>
<td>19</td>
<td>---</td>
<td>(3,4)</td>
</tr>
</tbody>
</table>
• `#pragma max_interleaving 1`
  The compiler restricts the annotated (inner) loop to be invoked only once per outer loop iteration. That is, all iterations of the inner loop travel the pipeline before the next invocation of the inner loop can occur.

• `#pragma max_interleaving 0`
  The compiler allows the pipeline to contain a number simultaneous invocations of the inner loop equal to the loop initiation interval (II) of the inner loop. For example, an inner loop with an II of 2 can have iterations from two invocations in the pipeline at a time.

  This behavior is the default behavior for the compiler if you do not specify the `max_interleaving` pragma.

In the following code snippet, the compiler restricts the pipelined execution of the `i` loop. A new invocation of the `i` loop corresponds only to subsequent iteration of the `j` loop.

```c
// Loop j is pipelined with ii=1
for (int j = 0; j < M; j++) {
  int a[N];
  // Loop i is pipelined with ii=2
  #pragma max_interleaving 1
  for (int i = 1; i < N; i++) {
    a[i] = foo(i)
  }
}
```

For another example of the effects of using the `max_interleaving` pragma, refer to the following tutorial:

<quartus_installdir>/hls/examples/tutorials/loop_controls/max_interleaving

### 6.9. Loop Fusion

Loop fusion is a compiler transformation in which two adjacent loops are merged into a single loop over the same index range. This transformation is typically applied to reduce loop overhead and improve run-time performance.

The following example shows the effects of fusing loops in a simple case:

<table>
<thead>
<tr>
<th>Unfused Loops</th>
<th>Fused Loop</th>
</tr>
</thead>
</table>
| for (j = 0; j < 300; j++)
  a[j] = a[j] + 3j
  b[k] = b[k] + 4j | for (f = 0; f < 300; f++) {
  a[f] = a[f] + 3f
  b[f] = b[f] + 4f |

The following example shows the effects of fusing loops over a concatenated loop-index range:

<table>
<thead>
<tr>
<th>Unfused Loops</th>
<th>Fused Loop</th>
</tr>
</thead>
</table>
| for (j = 0; j < 300; j++)
  a[j] = a[j] + 3j
  b[k] = b[k] + 4j | for (jk = 0; jk < 600; jk++) {
  if (jk < 300) {
    int j = jk;
    a[j] = a[j] + 3;
  } else
    |

Send Feedback
Unfused Loops

```c
int k = jk - 300;
if (k; < 300)
    b[k] = b[k] + 4;
}
```

Fused Loop

```c
int k = jk - 300;
if (k; < 300)
    b[k] = b[k] + 4;
}
```

Loop control structures represent a significant overhead. By fusing two loops, the number of control structures needed for the loops is reduced from two to one, reducing this overhead. The main goal of reducing the number of control structures is to free FPGA area for your design while still maintaining (ideally increasing) component throughput.

Fusing outer loops introduces parallelism where there was previously none. Combining bodies of two adjacent loops \( L_j \) and \( L_k \) forms a single loop \( L_f \) with a loop body that spans the bodies of \( L_j \) and \( L_k \). This combined loop body creates an opportunity for operations that were serialized across a given iteration of \( L_j \) and \( L_k \) to execute in parallel. In effect, the two loops now execute in lockstep as a single loop, which provides latency improvements.

If inner loops are fused, parallelism is already achieved by pipelined execution of the outer loop iteration. In these cases, the parallelism effect of loop fusion is diminished.

### Fusion Criteria

The compiler considers the fusion of two loops \( (L_j \) and \( L_k \)) to be valid if the loops meet the following criteria:

- The loops must be adjacent.
  - That is, you cannot have a statement \( S_i \) with side-effects such that \( S_i \) executes after \( L_j \) and before \( L_k \).
- Each loop must have a single-entry point and a single exit point.
  - For example, loops that contain `break` statements are not considered for fusion.
- The loops must have no negative-distance dependencies.
  - That is, for loops \( L_j \) and \( L_k \) where \( L_j \) is defined before \( L_k \), iteration \( m \) of loop \( L_k \) does not depend on values calculated in iteration \( m+n \) (where \( n>0 \)) of loop \( L_j \).

### Automatic Loop Fusion

The Intel HLS Compiler fuses adjacent loops with equal trip counts automatically if the compiler analysis of your component determines that fusing the loops is profitable.

Examples of where fusing loops is a valid transformation (based on the earlier criteria) but are not considered profitable by the compiler include the following situations:

- One of the two loops, but not both, is annotated with the `ivdep` pragma.
- One of the two loops, but not both, contains stall-free logic.

The Loop Analysis Report in the High-Level Design Reports indicates when loops were fused.

In addition to automatic loop fusion, the Intel HLS Compiler provides two pragmas to help you control when loops are fused:
• **nofusion pragma**
  Annotate loops with this pragma to request that the compiler not fuse the annotated loop.

• **loop_fuse pragma**
  Override the compiler profitability analysis and fuse adjacent loops provide that it is safe.

Use the the `loop_fuse` pragma to tell the compiler to consider fusing adjacent loops with different trip counts.

### 6.9.1. Loop Fusion Control (`loop_fuse`Pragma)

Use the `loop_fuse` pragma to tell the compiler to try to fuse two adjacent loops without affecting the functionality of either loop, overriding the compiler profitability analysis of fusing the loops.

Fusing adjacent loops helps reduce the amount of loop control overhead in your component, which helps reduce the FPGA area used and can increase the performance by executing both loops as one (fused) loop.

Apply the `loop_fuse` pragma to a block of code to indicate that loops in the code block should be considered for fusing as follows:

```c
#pragma loop_fuse [depth(N)] [independent]
{
  ...
}
```

By default, only adjacent top-level (not nested) loops are considered for fusing. Use the `depth(N)` clause of the pragma to indicate the number of nesting depths the compiler should consider when fusing adjacent loops. Specifying `depth(1)` is equivalent to indicating that only adjacent top-level loops should be considered for fusing.

```c
#pragma loop_fuse
// can also be
// #pragma loop_fuse depth(1)
{
  L1: for(...) {}
  L2: for(...) {
    L3: for(...) {}
    L4: for(...) {
      L5: for(...) {}
      L6: for(...) {}
    }
  }
}
```

By default (or `depth(1)`), only loops L1 and L2 are initially considered for fusing.

```c
#pragma loop_fuse depth(2)
{
  L1: for(...) {}
  L2: for(...) {
    L3: for(...) {}
    L4: for(...) {
      L5: for(...) {}
      L6: for(...) {}
    }
  }
}
```

With `depth(2)`, the following loop pairs are initially considered for fusing:

```c
#pragma loop_fuse depth(3)
{
  L1: for(...) {}
  L2: for(...) {
    L3: for(...) {}
    L4: for(...) {
      L5: for(...) {}
    }
    L6: for(...) {}
  }
}
```

With `depth(3)`, the following loop pairs are initially considered for fusing:

```c
#pragma loop_fuse depth(3)
{
  L1: for(...) {}
  L2: for(...) {
    L3: for(...) {}
    L4: for(...) {
      L5: for(...) {}
    }
    L6: for(...) {}
  }
}
```
The compiler automatically considers fusing adjacent loops with equal trip counts when the loops meet the criteria. You can also use the `loop_fuse` pragma to tell the compiler to consider fusing adjacent loops with different trip counts.

With the `loop_fuse` pragma applied to a block of code, the compiler always tries to fuse adjacent loops (with equal or different trip counts) in the block whenever the compiler determines that it is safe to fuse the loops. Two loops are considered safe to merge if they meet the fusion criteria described in Fusion Criteria section of Loop Fusion on page 71.

The following example shows the effects of fusing loops with unequal trip counts:

```c
#pragma loop_fuse
{
    for (int i = 0; i < N; i++) {
        // Loop Body 1
    }
    for (int j = 0; j < M; j++) {
        // Loop Body 2
    }
}
```

A fused loop can itself be considered for fusing with other loops. For example, in the following code L1 and L2 are initially considered for fusing. That resulting fused loop can then be considered for fusing with L4.

```c
#pragma loop_fuse
{
    L1 for(...) {}
    L2 for(...) {
        L3 for(...) {}
    }
    L4 for(...) {}
}
```

Use the `independent` option to override the dependency safety checks. If you specify the `independent` option, you are guaranteeing to the compiler that fusing pairs of loops affected by the `loop_fuse` pragma is safe. That is, there are no negative-distance dependencies in the fused loop. If it is not safe, you might get functional errors in your component.

**Function Calls In `loop_fuse` Code Blocks**

If a function call occurs in a code block annotated with the `loop_fuse` pragma and inlining that function call contains a loop, the resulting loop can be a candidate for loop fusion.

**Nested depth (N) Clauses**

When you nest `loop_fuse` pragmas, you might create overlapping sets of candidates loops.
Consider the following example:

```c
#pragma loop_fuse depth(2) independent
{ 
  L1: for(...) {} 
  L2: for(...) {
      #pragma loop_fuse depth(2)
      { 
        L3: for(...) {}
        L4: for(...) {
          L5: for(...) {}
          L6: for(...) {}
        }
      }
  }
}
```

In this example, the compiler considers the following loop pairs for fusion: L1/L2, L3/L4, and L5/L6. In addition, the compiler overrides the compiler negative-distance dependency analysis of the following loops pairs: L1/L2, L3/L4.

For another example of the effects of using the `loop_fuse` pragma, refer to the following tutorial:

<quartus_installdir>/hls/examples/tutorials/best_practices/loop_fusion

### 6.9.2. Loop Fusion Exemption (nofusion pragma)

You can exempt a loop from being considered for fusing with an adjacent loop by annotating the loop with the `nofusion` pragma. This pragma prevents the annotated loop from being automatically fused or fused when it is subject to the `loop_fuse` pragma.

Specify the `nofusion` pragma as follows:

```c
#pragma nofusion
for (...) {
  loop body
}
```

Applying the `nofusion` pragma to one of the loops in a pair prevents the loops from being fused.

For example, the following code samples have the same effect. If one loop in a pair is annotated with the `nofusion` pragma, the other loop has no other loop to fuse with.

```c
#pragma nofusion
L1: for (int j=0; j < N; ++j){
  data[j] += Q;
}
L2: for (int i = 0; i < N; ++i) {
  output[i] = Q * data[i];
}
```

For another example of the effects of using the `nofusion` pragma, refer to the following tutorial:

<quartus_installdir>/hls/examples/tutorials/best_practices/loop_fusion
7. Component Concurrency

The Intel HLS Compiler assumes that you want a fully pipelined data path in your component. In the C++ implementation, think of a fully pipelined data path as calling a function multiple times before the first call has returned (see also Figure 11 on page 60 and Intel HLS Compiler Hardware Model on page 14). That is, multiple invocations of a function, at different points in their executions, are executing at the same time. The Intel HLS Compiler might not be able to deliver a component with a component initiation interval (II) of 1, or even any pipelining if there are data or memory dependencies between invocations of the component.

The Intel HLS Compiler provides you with the hls_max_concurrency component attribute to help you control the maximum concurrency of your component.

7.1. Serial Equivalence within a Memory Space or I/O

Within a single memory space or I/O (stream read/write, Avalon-MM interface read/write, or component invocation input and return), every invocation of the component (that is, every cycle where the start signal is asserted and the component holds the busy signal low) on the component invocation interface behaves as though the previous invocation was fully executed.

When visualizing a single shared memory space, think of multiple function calls as executing sequentially, one after another. This way, when the component asserts the done signal, the results of a component invocation in hardware are guaranteed to be visible to both the next component invocation and the external system.

The HLS compiler takes advantage of pipeline parallelism to execute component invocations and loop iterations in parallel if the associated dependencies allow for parallel execution. Because the HLS compiler generates hardware that keeps track of dependencies across component invocations, it can support pipeline parallelism while guaranteeing serial equivalence across memory spaces. Ordering between independent I/O instructions is not guaranteed.

7.2. Concurrency Control (hls_max_concurrency Attribute)

You can use the hls_max_concurrency component attribute to increase or limit the maximum concurrency of your component. The concurrency of a component is the number of invocations of the component that can be in progress at one time. By default, the Intel HLS Compiler tries to maximize concurrency so that the component runs at peak throughput.
You can control the maximum concurrency of your component by adding the `hls_max_concurrency` component attribute immediately before you declare your component, as shown in the following example:

```c
#include "HLS/hls.h"

hls_max_concurrency(3)

component void foo ( /* arguments */ )
{
    // Component code
}
```

The optimizations caused by using this attribute might cause component memory configuration changes to meet the set concurrency requirements. Use memory attributes to control the geometry of your component memory configuration.

The Intel HLS Compiler restricts the maximum component concurrency to 1 in the following cases:

- **You have a component memory system.**
  
  At the component level, the Intel HLS compiler does not automatically create private copies of component memory to increase the throughput. If your component invocation uses a non-static component memory system, the next invocation cannot start until the previous invocation has finished all its accesses to and from that component memory.

  This limitation is shown in the Loop Analysis report as load-store dependencies on the component memory.

  Adding the `hls_max_concurrency(N)` attribute to the component creates private copies of the component memory so that you can have multiple pipelined invocations of your component in progress at the same time. To create as many private copies as necessary for maximal performance, use `hls_max_concurrency(0)`.

  For finer-grained control of which component memories to create private copies of, use the `hls_private_copies` memory attribute. For details, see `hls_private_copies` Memory Attribute.

- **The compiler determines that reducing concurrency saves FPGA area.**

  In some cases, the compiler reduces concurrency to save FPGA area. In these cases, the `hls_max_concurrency(N)` component attribute can increase the concurrency from 1.

  The Loop Analysis report displays the concurrency for a function in the Details pane of the report when you click the function marked with **(Component invocation)** in the Loop Analysis pane. If your design concurrency is limited, the Details pane shows a line like the following line:

  ```
  Maximum concurrent iterations: 1 is the default for component invocation loop.
  ```

  The `hls_max_concurrency` attribute can also accept a value of 0. When this attribute is set to 0, the component should be able to accept new invocations as soon as the downstream datapath frees up. Use this value only when you see loop initiation interval (II) issues in your component because using this attribute can increase the...
component area. You can find loop II issues by examining the Loop Viewer in the High-Level Design Reports or looking for extra bubbles that are visible in a simulation waveform.

You can also control the concurrency of loops in components with the `max_concurrency(N)` pragma. For more information about the `max_concurrency(N)` pragma, see Loop Concurrency (max_concurrency Pragma) on page 66.

### 7.3. Component Pipelining Control (hls_disable_component_pipelining Attribute)

If running concurrent invocations of your component does not improve throughput, or if you do not intend to invoke your component repeatedly, avoid extra FPGA area utilization by using the `hls_disable_component_pipelining` component attribute.

When you specify the `hls_disable_component_pipelining`, the Intel HLS Compiler generates a simpler, serialized datapath for your component.

You apply the attribute as shown in the following example:

```c
#include "HLS/hls.h"

hls_disable_component_pipelining
component void baz ( /* arguments */ ){
    // component code
}
```

You can also disable pipelining the datapath of a loop in your component with the `disable_loop_pipelining` pragma. For more information about this pragma see Loop Pipelining Control (disable_loop_pipeliningPragma) on page 68.

Review the Loop Analysis report in the High-Level Design Reports (report.html) to see component invocations and loops with pipelining disabled:
The Intel HLS Compiler Pro Edition supports a range of FPGA-optimized arbitrary-precision data types that are defined in header files that you can include in your designs.

Some of these header files are based on the Algorithmic C (AC) data types that Mentor Graphics* provides under the Apache license. For more information about the Algorithmic C data types, refer to Mentor Graphics Algorithmic C (AC) Datatypes, which is available as a part of your Intel HLS Compiler installation: <quartus_installdir>/hls/include/ref/ac_datatypes_ref.pdf.

The Intel HLS Compiler also supports arbitrary-precision IEEE 754 compliant floating point data types that is not based on the AC data types.

The Intel HLS Compiler supports the following arbitrary precision data types:

Table 19. Arbitrary Precision Data Types Supported by the Intel HLS Compiler Pro Edition

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Intel Header File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ac_int</td>
<td>HLS/ac_int.h</td>
<td>Arbitrary-width integer support</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To learn more, review the following tutorials:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• &lt;quartus_installdir&gt;/hls/examples/tutorials/ac_datatypes/ac_int_basic_ops</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• &lt;quartus_installdir&gt;/hls/examples/tutorials/ac_datatypes/ac_int_overflow</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• &lt;quartus_installdir&gt;/hls/examples/tutorials/best_practices/struct_interfaces</td>
</tr>
<tr>
<td>ac_fixed</td>
<td>HLS/ac_fixed.h</td>
<td>Arbitrary-precision fixed-point number support</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To learn more, review the tutorial:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• &lt;quartus_installdir&gt;/hls/examples/tutorials/ac_datatypes/ac_fixed_constructor</td>
</tr>
<tr>
<td></td>
<td>HLS/ac_fixed_math.h</td>
<td>Support for some nonstandard math functions for arbitrary-precision fixed-point data types</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To learn more, review the tutorial:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• &lt;quartus_installdir&gt;/hls/examples/tutorials/ac_datatypes/ac_fixed_math_library</td>
</tr>
<tr>
<td>ac_complex</td>
<td>HLS/ac_complex.h</td>
<td>Complex number support</td>
</tr>
<tr>
<td>hls_float</td>
<td>HLS/hls_float.h</td>
<td>Arbitrary-precision floating-point number support</td>
</tr>
<tr>
<td></td>
<td>HLS/hls_float_math.h</td>
<td>Support for commonly used exponential, logarithmic, power, and trigonometric functions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To learn more, review the following tutorials:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• &lt;quartus_installdir&gt;/hls/examples/tutorials/hls_float/1_reduced_double</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• &lt;quartus_installdir&gt;/hls/examples/tutorials/hls_float/2_explicit_arithmetic</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• &lt;quartus_installdir&gt;/hls/examples/tutorials/hls_float/3_conversions</td>
</tr>
</tbody>
</table>
The Intel HLS Compiler also supports some nonstandard math functions for the following data types when you include an additional header file:

- **ac_fixed data type**
  
  Include the HLS/ac_fixed_math.h header file

- **hls_float data type**
  
  Include the HLS/hls_float_math.h header file

**Advantages of Arbitrary Precision Data Types**

The arbitrary precision data types have the following advantages over using standard C/C++ data types in your components:

- You can achieve narrower data paths and processing elements for various operations in the circuit.
- The data types ensure that all operations are carried out in a size guaranteed not to lose any data. However, you can still lose data if you store data into a location where the data type is too narrow.

**Limitations of AC Data Types**

The AC data types have the following limitations:

- Multipliers are limited to generating 512-bit results.
- Dividers for ac_int data types are limited to a maximum of 128 bit unsigned or 127 bit signed.
- Dividers for ac_fixed data types are limited to a maximum of 64 bits (unsigned or signed).
- The FPGA-optimized header files provided by the Intel HLS Compiler are not compatible with GCC or MSVC. When you use the Intel HLS Compiler header files, you cannot use GCC or MSVC to compile your testbench. Both your component and testbench must be compiled with the Intel HLS Compiler.
  
  To compile AC data types with GCC or MSVC, use the reference AC data types headers also provided with the Intel HLS Compiler. For details, see AC Data Types and Native Compilers on page 90.

**Limitations of the Intel HLS Compiler Arbitrary Precision Floating Point Data Type**

The hls_float data type has the following limitations:

- Casting an hls_float datatype to an unsigned integer generates signed integer cast. Ensure that the hls_float value is cast into a signed integer that is large enough to contain the hls_float value being casted (for example, by using an ac_int datatype).
- Floating point optimization into constants performed for float and double data types is not done for the hls_float data type.
- A limited set of math functions is supported. For details, see Operators and Return Types Supported by the hls_float Data Type on page 92.
• The hls_float header files provided by the Intel HLS Compiler are not compatible with GCC or MSVC. When you use the Intel HLS Compiler header files, you cannot use GCC or MSVC to compile your testbench. Both your component and testbench must be compiled with the Intel HLS Compiler.

• The high-level design reports do not show bit widths for the hls_float data type.

• Constant initialization works only with the round-towards-zero (RZERO) rounding mode.

Related Information
AC Datatypes at HLSLibs

8.1. Declaring ac_int Data Types

The HLS compiler package includes an ac_int.h header file to provide arbitrary precision integer support in your component.

1. Include the ac_int.h header file in your component in the following manner:

```c
#ifdef __INTELFPGA_COMPILER__
#include "HLS/ac_int.h"
#else
#include "ref/ac_int.h"
#endif
```

2. After you include the header file, declare your ac_int variables in one of the following ways:

   — Template-based declaration
     - ac_int<N, true> var_name; //Signed N bit integer
     - ac_int<N, false> var_name; //Unsigned N bit integer
   — Predefined types up to 63 bits
     - intN var_name; //Signed N bit integer
     - uintN var_name; //Unsigned N bit integer

Where N is the total length of the integer in bits.
Restriction: If you want to initialize an `ac_int` variable to a value larger than 64 bits, you must use the `bit_fill` or `bit_fill_hex` utility function. For details see "2.3.14 Methods to Fill Bits" in Mentor Graphics Algorithmic C (AC) Datatypes, which is available as `<quartus_installdir>/hls/include/ref/ac_datatypes_ref.pdf`.

The following code example shows the use of the `bit_fill` or `bit_fill_hex` utility functions:

```c
typedef ac_int<80,false> i80_t;

i80_t x;

x.bit_fill_hex("a9876543210fedcba987"); // member function
x = ac::bit_fill_hex<i80_t>("a9876543210fedcba987"); // global function

int vec[] = { 0xa987, 0x6543210f, 0xedcba987 };
x.bit_fill(vec); // member function
x = bit_fill<i80_t>(vec); // global function
// inlining the constant array
x.bit_fill( (int [3]) { 0xa987,0x6543210f,0xedcba987 } ); // member function
x = bit_fill<i80_t>( (int [3]) { 0xa987,0x6543210f,0xedcba987 } ); // global function
```

For a list of supported operators and their return types, see "Chapter 2: Arbitrary-Length Bit-Accurate Integer and Fixed-Point Datatypes" in Mentor Graphics Algorithmic C (AC) Datatypes, which is available in the following file: `<quartus_installdir>/hls/include/ref/ac_datatypes_ref.pdf`.

### 8.1.1. Important Usage Information on the `ac_int` Data Type

The `ac_int` datatype has a large number of API calls that are documented in Mentor Graphics Algorithmic C (AC) Datatypes, which is available in your Intel HLS Compiler installation as the following file: `<quartus_installdir>/hls/include/ref/ac_datatypes_ref.pdf`. For more information on AC datatypes, refer to

The `ac_int` datatype automatically increases the size of the result of the operation to guarantee that the intermediate operations never overflow. However, the HLS compiler automatically truncates or extends the result to the size of the specified destination container, so ensure that the storage variable for your computation is large enough.

The HLS compiler installation package includes a number of examples in the tutorials. Refer to the tutorials in `<quartus_installdir>/hls/example/tutorials/ac_datatypes` for some of the recommended practices.

### 8.1.2. Integer Promotion and `ac_int` Data Types

The rules of integer promotion when you use `ac_int` data types are different from the rules of integer promotion for standard C/C++ rules. Your component design should account for these differing rules.

Depending on the data type of the operands, integer promotion is carried out differently:
- Both operands are standard integer types (int, short, long, unsigned char, or signed char):
  
  If both operands are of standard integer type (for example char or short) operations, integers can be promoted following the C/C++ standard. That is, the operation is carried out in the data type and size of the largest operand, but at least 32 bits. The expression returns the result in the larger data type.

- Both operands are ac_int data types:
  
  If both operands are ac_int data types, operations are carried out in the smallest ac_int data type needed to contain all values. For example, the multiplication of two 8-bit ac_int values is carried out as an 16-bit operation. The expression returns the result in that type.

- One operand is a standard integer type and one operand is an ac_int type:
  
  If the expression has one standard data type and one ac_int type, the rules for ac_int data type promotion apply. The resulting expression type is always an ac_int data type. For example, if you add a short data type and an ac_int<16> data type, the resulting data type is ac_int<17>.

### Data Width of Operations

C++ compilers typically automatically promote narrow integer types such as char and short to 32-bit types (int) for arithmetic operations such as addition, multiplication, division, and bit-shifts. To adhere to the C++ language specification and be consistent with other C++ compilers, the Intel HLS Compiler might use larger operations than you might expect when dealing with native types.

If you need better control over the size of arithmetic operations, use the ac_int datatype.

Consider the following example:

```cpp
cOMPONENT int singlestep_nativeSigned(char a, char b, char c)
{
  return a * b / c;
}
```

The System Viewer (part of the High-Level Design Reports) shows that the divide operation becomes a 32-bit operation:
The Intel HLS Compiler narrows operations automatically whenever it can, depending on the context of the operation. In this case, the compiler automatically narrowed the multiplication operations to 16 bits. To confirm the width of an operation, review your design in the System Viewer.

If you use `ac_int` data types for the parameters and the return data, the Intel HLS Compiler makes the divide operation narrower.

```c
#include <HLS/ac_int.h>
component int32 singlestep_acint_signed(int8 a, int8 b, int8 c)
{
    return a * b / c;
}
```

With the component variables all defined as `ac_int` data types, the System Viewer now shows that the divide operation is reduced to being 17-bits wide:
**Literals in Operations**

In C/C++, literals are by default an `int` data type, so when you use a literal without any casting, the expression type is always at least 32 bits. For example, if you have code like following code snippet, the comparison is carried out in 32 bits:

```c
ac_int<5, true> ap;
...
if (ap < 4) {
...
```

If the operands are signed differently and the unsigned type is at least as large as the signed type, the operation is carried out as an unsigned operations. Otherwise, the unsigned operand is converted to a signed operand.

For example, if you have code like the following snippet, the `-1` value expands to a 32-bit negative number (0xffffffff) while the `uint3` value is a positive 32-bit number 7 (0x00000007):

```c
uint3 x = 7;
if (x != -1) {
    // FAIL
}
To keep the compiler from expanding the literal to 32-bits, cast the literal to the same `ac_int` data type. For example:

```c
uint3 x = 7;
if (x != (unit3)-1) { // SUCCEED
    
}
```

### 8.1.3. Debugging Your Use of the `ac_int` Data Type

The "HLS/ac_int.h" header file provides you with tools to help check `ac_int` operations and assignments for overflow in your component when you run an x86 emulation of your component: `DEBUG_AC_INT_WARNING` and `DEBUG_AC_INT_ERROR`.

When you use the `DEBUG_AC_INT_WARNING` and `DEBUG_AC_INT_ERROR` macros, you cannot declare `constexpr ac_int` variables or `constexpr ac_int` arrays.

<table>
<thead>
<tr>
<th>Tool</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEBUG_AC_INT_WARNING</td>
<td>Emits a warning for each detected overflow.</td>
</tr>
<tr>
<td>DEBUG_AC_INT_ERROR</td>
<td>Emits a message for the first overflow that is detected and then exits the component with an error.</td>
</tr>
</tbody>
</table>

After you use these tools to determine that your component has overflows, run the `gdb` debugger on your component to run the program again and step through the program to see where the overflows happen.

Review the `ac_int_overflow` tutorial in `<quartus_installdir>/hls/example/tutorials/ac_datatypes` to learn more.

### 8.2. Declaring `ac_fixed` Data Types

The HLS compiler package includes an `ac_fixed.h` header file for arbitrary precision fixed-point support.

1. Include the `ac_fixed.h` header file in your component in the following manner:

   ```c
   #ifdef __INTELFPGA_COMPILER__
   #include "HLS/ac_fixed.h"
   #else
   #include "ref/ac_fixed.h"
   #endif
   ```

2. After you include the header file, declare your `ac_fixed` variables as follows:

   - `ac_fixed<N, I, true, Q, O> var_name; //Signed fixed-point number`
   - `ac_fixed<N, I, false, Q, O> var_name; //Unsigned fixed-point number`

   Where the template attributes are defined as follows:

   - N: The total length of the fixed-point number in bits.
The number of bits used to represent the integer value of the fixed-point number.

The difference of $N - I$ determines how many bits represent the fractional part of the fixed-point number.

The quantization mode that determines how to handle values where the generated precision (number of decimal places) exceeds the number of bits available in the variable to represent the fractional part of the number.

For a list of quantization modes and their descriptions, see "2.1. Quantization and Overflow" in Mentor Graphics Algorithmic C (AC) Datatypes, which is available in the following file:

<quartus_installdir>/hls/include/ref/ac_datatypes_ref.pdf.

The overflow mode that determines how to handle values where the generated value has more bits than the number of bits available in the variable.

For a list of overflow modes and their descriptions, see "2.1. Quantization and Overflow" in Mentor Graphics Algorithmic C (AC) Datatypes, which is available in the following file:

<quartus_installdir>/hls/include/ref/ac_datatypes_ref.pdf.

For a list of supported operators and their return types, see "Chapter 2: Arbitrary-Length Bit-Accurate Integer and Fixed-Point Datatypes" in Mentor Graphics Algorithmic C (AC) Datatypes, which is available in the following file:

<quartus_installdir>/hls/include/ref/ac_datatypes_ref.pdf.

Additional math functions are supported by the HLS/ac_fixed_math.h header file. For details, see Math Functions Provided by the ac_fixed_math.h Header File on page 210.

Important: Due to the differences in the internal math implementations, the results from operations with ac_fixed data types might be different between simulation and emulation. The maximum difference will be within a few ULPs.

### 8.2.1. Arbitrary Precision Fixed-Point Literals in Operations

When you use the ac_fixed datatype to declare literals in your component, the compiler expands the literal, regardless of the size you set for the ac_fixed literal.

To keep the compiler from expanding the literal, cast the literal to an ac_fixed data type.

For example, the literal in the following code snippet is treated as a double data type:

Using Fixed12 = ac_fixed<12, 6, true>;

```c
component Fixed12 threshold_vanilla(Fixed12 a){
    if (a < 12.7) {
        return a;
    } else {
        return 12.7;
    }
}
```

**Important:** Due to the differences in the internal math implementations, the results from operations with ac_fixed data types might be different between simulation and emulation. The maximum difference will be within a few ULPs.
In this code snippet, the 12-bit `ac_fixed` value is compared with 12.7. The literal 12.7 is treated as a `double`, with value 12.699999... This literal is converted to a fixed-point to be compared with `a`, but because the literal is unconstrained, the compiler chooses a large fixed-point value and this can result in a large comparison:

```
Using Fixed12 = ac_fixed<12, 6, true>;
component Fixed12 threshold_cast(Fixed12 a)
if (a < (Fixed12)12.7) {
```

You can force a smaller comparison by casting the literal 12.7 to an appropriately-sized `ac_fixed` data type:
return a;
} else {
    return 12.7;
}

For more details about the impact of choosing single precision literals and functions over their double precision counterparts, review the following tutorial:

<quartus_installdir>/hls/examples/tutorials/best_practices/single_vs_double_precision_math

### 8.3. Declaring ac_complex Data Types

The HLS compiler package includes an ac_complex.h header file for complex number support.
1. Include the `ac_complex.h` header file in your component in the following manner:

```c
#ifdef __INTELFPGA_COMPILER__
#include "HLS/ac_complex.h"
#else
#include "ref/ac_complex.h"
#endif
```

2. After you include the header file, declare your `ac_complex` variables according to the data type of your complex number. The underlying data type can be `ac_int`, `ac_fixed`, `hls_float`, and standard C integer or floating-point data types.

For a list of supported operators and their return types, see "4. Complex Datatype" in Mentor Graphics Algorithmic C (AC) Datatypes, which is available in the following file:

```plaintext
<quartus_installdir>/hls/include/ref/ac_datatypes_ref.pdf
```

### 8.4. AC Data Types and Native Compilers

The reference version of the Mentor Graphics Algorithmic C (AC) data types is also provided with the Intel HLS Compiler. Do not use these reference header files in your component if you want to compile your component with an FPGA target.

Use the reference header files for AC data types to confirm functional correctness in your component when you are compiling your component with native compilers (g++ or MSVC).

If you use the reference header files and compile your component to an FPGA target, your component can compile successfully but your component QoR will be poor.

All of your code must use the same header files (either the reference header files or the FPGA-optimized header files). For example, your code cannot use the reference header files in your testbench and, at the same time, use the FPGA-optimized header file in your component code.

The following reference header files are provided with the Intel HLS Compiler:

<table>
<thead>
<tr>
<th>AC data type</th>
<th>Reference Header File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ac_int</code></td>
<td>ref/ac_int.h</td>
<td>Arbitrary width integer support</td>
</tr>
<tr>
<td><code>ac_fixed</code></td>
<td>ref/ac_fixed.h</td>
<td>Arbitrary precision fixed-point number support</td>
</tr>
<tr>
<td><code>ac_complex</code></td>
<td>ref/ac_complex.h</td>
<td>Arbitrary precision complex number support</td>
</tr>
</tbody>
</table>

### 8.5. Declaring `hls_float` Data Types

The Intel HLS Compiler Pro Edition includes the `hls_float.h` header file for arbitrary-precision floating-point number support. The floating-point representation for `hls_float` data types adopts the same IEEE 754 standard as native C++ `float` and `double` types.

The `hls_float.h` header file does not work with native compilers (g++ or MSVC).

An `hls_float` variable carries an explicit sign bit and an arbitrary number of bits for the exponent and mantissa.
Due to the differences in the internal math implementations and rounding errors, the results from hls_float operations might not always be bit-accurate to those produced by C++ native floating-point types with the same exponent and mantissa bit widths.

1. Include the hls_float.h header file in your component in the following manner:

   ```c
   #include "HLS/hls_float.h"
   ```

2. After you include the header file, declare your hls_float variables as follows:

   ```c
   hls_float<exponent_width, mantissa_width>[rounding_mode]>
   ```

   The hls_float.h header file also provides aliases in the ihc namespace so that you can declare ihc::bfloat16 and ihc::bfloat19 data types directly.

   Where the template attributes are defined as follows:

   - `exponent_width`, `mantissa_width`
     The bit-width of the exponent and mantissa of the floating-point variable.

   The hls_float data type supports the following `exponent_width, mantissa_width` combinations:

<table>
<thead>
<tr>
<th><code>exponent_width, mantissa_width</code></th>
</tr>
</thead>
<tbody>
<tr>
<td>5, 10</td>
</tr>
<tr>
<td>8, 7</td>
</tr>
<tr>
<td>8, 10</td>
</tr>
<tr>
<td>8, 17</td>
</tr>
<tr>
<td>8, 23</td>
</tr>
<tr>
<td>8, 26</td>
</tr>
<tr>
<td>10, 35</td>
</tr>
<tr>
<td>11, 44</td>
</tr>
<tr>
<td>11, 52</td>
</tr>
<tr>
<td>15, 63</td>
</tr>
</tbody>
</table>

   Some of these width combinations map to some commonly used floating-point formats:

<table>
<thead>
<tr>
<th>Floating-Point Format</th>
<th><code>exponent_width, mantissa_width</code> Setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>IEEE 754 half-precision (binary16)</td>
<td>5, 10</td>
</tr>
<tr>
<td>bfloat16(2)</td>
<td>8, 7</td>
</tr>
<tr>
<td>bfloat19(3)</td>
<td>8, 10</td>
</tr>
<tr>
<td>IEEE 754 single-precision (binary32)</td>
<td>8, 23</td>
</tr>
<tr>
<td>IEEE 754 double-precision (binary64)</td>
<td>11, 52</td>
</tr>
<tr>
<td>80-bit extended precision(4)</td>
<td>15, 63</td>
</tr>
</tbody>
</table>

(2) You can also declare this format directly as a bfloat16 data type.

(3) You can also declare this format directly as a bfloat19 data type.

(4) Not a bit-to-bit mapping.

80-bit extended precision has one explicit bit of fraction that is dropped when converting it to hls_float<15,63>.
Optional parameter to specify the IEEE 754 rounding mode used when converting between data types.

Set the rounding mode with one of the following values:

- `ihc::fp_config::FP_Round::RNE`
  - Round to nearest, tie to even
  - This rounding mode is more accurate (0.5 ULP), but requires more FPGA area.
- `ihc::fp_config::FP_Round::RZERO`
  - Round towards zero
  - This rounding mode is less accurate (1 ULP) and requires less FPGA area.

If you do not set this parameter, the Intel HLS Compiler uses the `ihc::FP_Round::RNE` rounding mode.

The `hls_float` data type supports a limited set of math operations. For details, see Operators and Return Types Supported by the `hls_float` Data Type on page 92.

### 8.5.1. Operators and Return Types Supported by the `hls_float` Data Type

The `hls_float` data type supports all overloaded math operators and a limited set of the math functions provided by the Intel HLS Compiler Pro Edition. For some math operators, you can control the precision of the output by using templated versions of the functions.

**Important:** Due to the differences in the internal math implementations and rounding errors, the results from `hls_float` operations might not always be bit-accurate to those produced by C++ native floating-point types with the same exponent and mantissa bit widths. However, these results are validated against the infinitely accurate results.

**Supported Math Functions**

In addition to supporting all overloaded math operators, the Intel HLS Compiler supports the following additional math functions for the `hls_float` data type through the `HLS/hls_float_math.h` header file:

- Exponential and logarithmic functions(*):
  - `ln, log2, log10, ln(1+x)`
  - `e^x, 2^x, 10^x, e^{x-1}`

(*) Not supported for `hls_float<15,63>` precision variables.
• Advanced functions(*):
  — reciprocal
  — reciprocal_sqrt
  — sqrt(*)
  — cbrt (cube root)
  — hypot (hypotenuse)
• Power functions(*):
  — pow, powr, pown
• Trigonometric functions(*):
  — sin, cos, sincos
  — sinpi, cospi
  — asin, asinpi, acos, acospit, atan, atanpi, atan2

Conversion Rules

You can convert between different sizes of hls_float data types through assignment or by using the convert_to() function. For example,

```cpp
hls_float<8, 32> myFloat = ...;
hls_float<3, 18> myFloat2 = myFloat; // use rounding rules defined by hls_float type
hls_float<3, 18> myFloat3 = myFloat.convert_to<3, 18, ihc::fp_config::FP_Round::RZERO>(); // use rounding rules defined in convert_to() function call
```

To convert between native types (for example, float, double) and hls_float data types, assign to or from the types. Type conversion in an assignment occurs according to the rules in the Default Conversion Rules for hls_float Variables table that follows.

For two hls_float variables in a binary operation, the hls_float variable with the larger exponent bit-width is considered to be the "larger" variable. If the two variables have the same exponent bit width, the variable with the larger mantissa bit-width is considered to be the larger variable. The operands are then unified to the "larger" type before the binary operation occurs.

Native floating point data types and hls_float data types are converted to hls_float data types according to the rules in the Default Conversion Rules for hls_float Variables table that follows.

The Intel HLS Compiler also provides some operations that leave the precision of input types untouched and provide control over the output precision. For details, see Operations With Explicit Precision Controls on page 94.

Table 22. Default Conversion Rules for hls_float Variables

<table>
<thead>
<tr>
<th>Data Type</th>
<th>From hls_float To Data Type</th>
<th>From Data Type To hls_float</th>
</tr>
</thead>
<tbody>
<tr>
<td>hls_float with higher representable range</td>
<td>Keep exponent equivalent.</td>
<td>+-Inf if the source of the conversion is out of the representable range. Otherwise, keep exponent equivalent.</td>
</tr>
</tbody>
</table>

continued...
## Data Type Conversion

<table>
<thead>
<tr>
<th>Data Type</th>
<th>From hls_float To Data Type</th>
<th>From Data Type To hls_float</th>
</tr>
</thead>
<tbody>
<tr>
<td>float</td>
<td>The mantissa is rounded according to the rounding mode of the target hls_float (with the higher representable range).</td>
<td>The mantissa is rounded according to the rounding mode of the target hls_float (with the smaller representable range).</td>
</tr>
<tr>
<td>double</td>
<td>Convert original hls_float to hls_float&lt;8, 23&gt; with earlier hls_float rule, then bit-cast to float</td>
<td>Bit-cast float to hls_float&lt;8, 23&gt;, and then convert target hls_float precision using the hls_float to hls_float rules described earlier.</td>
</tr>
<tr>
<td>long double (emulation only) (Linux only)</td>
<td>Convert original hls_float to hls_float&lt;11, 52&gt; with earlier hls_float rule, then bit-cast to double</td>
<td>Bit-cast double to hls_float&lt;11, 52&gt;, and then convert target hls_float precision using the hls_float to hls_float rules described earlier.</td>
</tr>
<tr>
<td>long double (emulation only) (Windows only)</td>
<td>Same as double</td>
<td>Same as double</td>
</tr>
<tr>
<td>C++ native integer types</td>
<td>Truncate towards zero</td>
<td>Round to nearest, tie breaks to even. If the integer value is too large, the hls_float value saturates to plus infinity.</td>
</tr>
</tbody>
</table>

## Operations With Explicit Precision Controls

The Intel HLS Compiler provides the following operations that leave the precision of input hls_float-type variables untouched and let you control the output precision:

### Rounding Mode Control For hls_float to hls_float Conversions

**Syntax**

```cpp
convert_to<output_exponent_width, output_mantissa_width, rounding_mode>
```

**Description**

Use this method to override the rounding mode set for an hls_float variable when you are converting the variable to different precision.

By default, hls_float to hls_float conversions use the rounding mode that you specified when you declared the variable.

### Multiplication

**Syntax**

```cpp
ihc::hls_float< output_exponent_width, output_mantissa_width > ::mul <accuracy_setting>, [subnormal_setting]>(hls_float_a, hls_float_b)
```

Where the optional parameters are defined as follows:

- **subnormal_setting**: Optional parameter to specify whether input and output number are flushed to zero when carrying out basic binary operations explicitly.
Set this parameter with one of the following values:

- **ihc::fp_config::FP_Subnormal::ON**
  Input and output numbers in the subnormal range are preserved.
  The target FPGA device must have subnormal support,
  Subnormal support might require more FPGA area.
- **ihc::fp_config::FP_Subnormal::OFF**
  Input or output numbers in the subnormal range are flushed to zero.
- **ihc::fp_config::FP_Subnormal::AUTO**
  With this setting, the Intel HLS Compiler enables subnormal support only when it is directly supported by the target FPGA device and it does incur any extra FPGA area overhead.

If you do not set this parameter, the Intel HLS Compiler uses the **ihc::FP_Subnormal::AUTO** subnormal setting.

**accuracy_setting**

Optional parameter that influences trade-offs between the accuracy of the result due to different rounding decisions in the intermediary calculations and the FPGA area utilized by the generated hardware. Floating-point operations with less accurate results typically use fewer logic elements.

For example, a divider with a high accuracy might use 20% more FPGA area than divider with low accuracy. The low accuracy divider has a higher error bound [1 unit of least precision (ULP)] than a high accuracy divider (0.5 ULP).

Set this parameter with one of the following values:

- **ihc::fp_config::FP_Accuracy::LOW**
- **ihc::fp_config::FP_Accuracy::HIGH**

If you do not set this parameter, the Intel HLS Compiler uses the **ihc::fp_config::FP_Accuracy::HIGH** accuracy setting.

**Description**

This math function supplements the basic multiplication operation performed by the multiplication (*) operator.

Multiplies hls_float_a and float_b without changing the input types, and outputs an hls_float at the specified precision.
Addition/Subtraction/Division

Syntax

ihc::hls_float< output_exponent_width, output_mantissa_width > ::add <[optional parameters]> (hls_float_a, hls_float_b)

ihc::hls_float< output_exponent_width, output_mantissa_width > ::sub <[optional parameters]> (hls_float_a, hls_float_b)

ihc::hls_float< output_exponent_width, output_mantissa_width > ::div <[optional parameters]> (hls_float_a, hls_float_b)

Description

These math functions supplement the basic math operations performed by the addition/subtraction/division (+/−/) operators.

Adds/Subtracts/Divides hls_float_a and hls_float_b by first casting hls_float_a and hls_float_b to the specified hls_float precision. The operation and output are at the specified precision.

You can also specify the optional parameters that are the accuracy Setting and subnormal Setting parameters described earlier.

Comparison Operators

Comparison operators (> , < = , != , >= , <=) are subject to the conversion rules described earlier.

The == and != operators impose a bit-wise comparison of the casted values.

Comparisons with NaN always return false.

Additional hls_float Functions

The hls_float data type also has the following additional functions:

<p>| Table 23. |</p>
<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Getters and Setters</td>
<td></td>
</tr>
<tr>
<td>hls_float::get_exponent</td>
<td>Gets the exponent value of the hls_float variable.</td>
</tr>
<tr>
<td>hls_float::set_exponent</td>
<td></td>
</tr>
<tr>
<td>hls_float::get_mantissa</td>
<td></td>
</tr>
<tr>
<td>hls_float::set_mantissa</td>
<td></td>
</tr>
<tr>
<td>hls_float::get_sign</td>
<td>Gets the sign bit of the hls_float variable.</td>
</tr>
<tr>
<td>hls_float::set_sign</td>
<td></td>
</tr>
<tr>
<td>Special Constants</td>
<td></td>
</tr>
<tr>
<td>hls_float&lt;e,m&gt;::nan()</td>
<td>Constant used to assign the hls_float variable a value of NaN.</td>
</tr>
<tr>
<td>hls_float&lt;e,m&gt;::pos_inf()</td>
<td>Constant used to assign the hls_float variable a value of +\infty.</td>
</tr>
</tbody>
</table>
8.5.2. Additional Data Types Provided By hls_float.h

The Intel HLS Compiler hls_float.h header files provides some aliases in the ihc namespace for some hls_float data types that you can use instead of explicitly declaring an hls_float data type.

**The ihc::bfloat16 Data Type**

The bfloat16 data type is a 16-bit floating point number with an 8-bit exponent and a 7-bit mantissa (equivalent to declaring hls_float<8.7>).

On Intel Agilex devices, dot product operations that involve the bfloat16 (or hls_float<8.7>) data type are mapped to FP16 digital signaling blocks (DSPs). On other device families, dot product operations are mapped to adaptive logic modules (ALMs) and fixed-point 18-bit DSPs.

On all device families, all other math functions are mapped to ALMs and fixed-point 18-bit DSPs.

**The ihc::bfloat19 Data Type**

The bfloat19 data type is a 19-bit floating point number with an 8-bit exponent and a 10-bit mantissa (equivalent to declaring hls_float<8.10>).

On Intel Agilex devices, dot product operations that involve the bfloat19 (or hls_float<8.10>) data type are mapped to FP19 digital signaling blocks (DSPs). On other device families, dot product operations are mapped to adaptive logic modules (ALMs) and fixed-point 18-bit DSPs.

On all device families, all other math functions are mapped to ALMs and fixed-point 18-bit DSPs.
9. Component Target Frequency

You can specify component target frequency either in the i++ command by specifying the --clock option or by using the hls_scheduler_target_fmax_mhz component attribute. The component attribute takes priority over the command option.

For details about the --clock option, see Command Options Affecting Compiling on page 144.

For details about the hls_scheduler_target_fmax_mhz component attribute, see hls_scheduler_target_fmax_mhz Component Attribute on page 170.

The two options for setting target frequency are functionally equivalent except their scopes differ:

- The --clock option applies to all components compiled with the invocation of the i++ command that contains the --clock option.
- The hls_scheduler_target_fmax_mhz component attribute applies only to the component or task function that has the attribute.

To learn more about the attribute and how it interacts with the loop pragma, review the following tutorial:

<quartus_installdir>/hls/examples/tutorials/best_practices/set_component_target_fmax

If you use both the i++ command --clock option and the hls_scheduler_target_fmax_mhz component attribute, the component attribute takes priority. For example, you can compile the following code with the i++ ... --clock=300MHz command:

```c
component int test1(){
  ...
}

hls_scheduler_target_fmax_mhz(200)
component int test2(){
  ...
}
```

The compiler schedules component test1 at 300 MHz (from the command option) and component test2 at 200 MHz (from the component attribute).

**Important!** Setting the target $f_{\text{MAX}}$ determines the pipelining effort at the compilation stage. Compiling with Quartus Prime software reports the achievable $f_{\text{MAX}}$ value for your components. This value is often different from the value you specified.
You can lower the --clock value to reduce the latency of your design at the expense of reducing the \( f_{\text{MAX}} \) of your component.

### 9.1. Effects of Specifying Target II and Target \( f_{\text{MAX}} \)

Setting a target II (through `hls_component_ii` component attribute or `#pragma ii` loop pragma) and setting a target \( f_{\text{MAX}} \) (through the `hls_scheduler_target_fmax_mhz` component attribute) affects how the scheduler in the Intel HLS Compiler determines its efforts.

The following table summarizes the behavior of the scheduler in the Intel HLS Compiler.

<table>
<thead>
<tr>
<th>Set Target ( f_{\text{MAX}} )</th>
<th>Set Target II</th>
<th>Compiler Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>No</td>
<td>Yes</td>
<td>Best effort to achieve the II for the corresponding loop (may not achieve the best possible ( f_{\text{MAX}} )).</td>
</tr>
<tr>
<td>Yes</td>
<td>No</td>
<td>Best effort to achieve ( f_{\text{MAX}} ) specified (may not achieve the best possible II).</td>
</tr>
<tr>
<td>Yes</td>
<td>Yes</td>
<td>Best effort to achieve the ( f_{\text{MAX}} ) specified at the given II. The compiler errors out if it cannot achieve the requested II.</td>
</tr>
<tr>
<td>No</td>
<td>No</td>
<td>Use heuristic to achieve best ( f_{\text{MAX}}/II ) trade-off.</td>
</tr>
</tbody>
</table>

**Important:** If you are using an \( f_{\text{MAX}} \) target in the command line or a component attribute, specify `#pragma ii` for performance-critical loops in your design.
10. Systems of Tasks

Your component design might contain operations that you want to run asynchronously from the main flow of your component. The Intel HLS Compiler Pro Edition lets you define these asynchronous activities in task functions. These task functions, along with the component that invokes them, constitute a system of tasks.

The component keyword marks a single function and its subfunctions as a component. Within this component function, directly-called functions are in-lined while functions that use the systems of tasks API calls (ihc::launch and ihc::collect) generate hardware outside the component datapath and behave like an asynchronous call.

The function tagged with the component keyword marks the boundary of a system of tasks. Your external system can interact with all the interfaces that the component exposes.

Implementing your design as a system of tasks instead of a monolithic component can be useful in situations where expressing coarse-grained thread-level parallelism is needed. For example, a system of tasks is useful in the following situations:

- Improving the performance of operations like executing loops in parallel
- Reducing FPGA area utilization by sharing an expensive compute block with different parts of your component

### Table 24. Intel HLS Compiler System of Tasks Summary

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::launch</td>
<td>Marks a function as an Intel HLS Compiler task for hardware generation, and launches the task function asynchronously.</td>
</tr>
<tr>
<td>ihc::collect</td>
<td>Synchronizes the completion of the specified task function in the component.</td>
</tr>
<tr>
<td>ihc::stream</td>
<td>Allows streaming communication between different task functions.</td>
</tr>
<tr>
<td>ihc::launch_always_run</td>
<td>Launches a task function at component power-on or reset and continuously executes the function. <em>Recommendation:</em> Use the ihc_hls_set_component_wait_cycle with this function to keep your component and always-run task functions correctly coordinated.</td>
</tr>
</tbody>
</table>
Table 25. Intel HLS Compiler Systems of Tasks Streaming Interface Template Summary

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::stream</td>
<td>Streaming interface to the component or task function.</td>
</tr>
<tr>
<td>ihc::buffer</td>
<td>Specifies the capacity (in words) of the FIFO buffer on the input data that associates with the stream.</td>
</tr>
<tr>
<td>ihc::usesPackets</td>
<td>Exposes the startofpacket and endofpacket sideband signals on the stream interface.</td>
</tr>
</tbody>
</table>

Table 26. Intel HLS Compiler Streaming Input Interface stream Function APIs

<table>
<thead>
<tr>
<th>Function API</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T read()</td>
<td>Blocking read call to be used from within the component or task.</td>
</tr>
<tr>
<td>T read(bool &amp; sop, bool &amp; eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Blocking read with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>T tryRead(bool &amp; success)</td>
<td>Non-blocking read call to be used from within the component or task. The success bool is set to true if the read was valid.</td>
</tr>
<tr>
<td>T tryRead(bool &amp; success, bool &amp; sop, bool &amp; eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Non-blocking read with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>void write(T data)</td>
<td>Blocking write call from the component or task.</td>
</tr>
<tr>
<td>void write(T data, bool sop, bool eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Blocking write with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>bool tryWrite(T data)</td>
<td>Non-blocking write call from the component or task. The return value represents whether the write was successful.</td>
</tr>
<tr>
<td>bool tryWrite(T data, bool sop, bool eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Non-blocking write with out-of-band startofpacket and endofpacket signals. The return value represents whether the write was successful.</td>
</tr>
</tbody>
</table>

10.1. Task Functions

The Intel HLS Compiler Pro Edition implements task functions in a way similar to HLS component functions, but with some additional constraints.

Task Function Interfaces

You can get data in and out of task functions in the following ways:

- Task function arguments and return values
- Explicitly-declared Avalon Streaming (ST) interfaces
- Explicitly-declared Avalon Memory-Mapped (MM) Host interfaces

Task function interfaces are subject to the following restrictions:
You cannot use Avalon MM Host interfaces (ihc::mm_host) defined at global or file scope in a component or its tasks.

Pointer or reference data types cannot be passed to task functions as arguments unless they are a pointer or reference to an explicit Avalon MM Host interface.

For more details and example of task function interfaces, refer to the following tutorial:

<quartus_installdir>/hls/examples/tutorials/system_of_tasks/interfaces_sot

**Task Invocation Interface**

You can invoke a function as a task with the ihc::launch<>() and ihc::collect<>() API.

An ihc::launch<task_function>() call in your code produces a call interface for the specified task with the top-level component as the source and the task as the sink. The top-level component invokes the call interface of a task by asserting the start signal to start a new execution of the task.

An ihc::collect<task_function>() call in your code produces a return interface for the specified task with the task as the source and the top-level component as the sink. This return interface has a done signal that is asserted by the task when it has completed a single execution. When the top-level component invokes the return interface it consumes this done signal and retrieves any returned values.

In the High Level Design Reports (report.html), the ihc::launch and ihc::collect calls appear as blocking streaming write and streaming read operations.

**Scalar Parameters and Return Values**

Like HLS components, the scalar parameters and return value for an HLS task are implemented as conduits and the hand-shaking is implemented as a simple stall/valid handshake. The ihc::launch and ihc::collect calls connect directly to the HLS task function do and return streams.

In the High Level Design Report (report.html), the ihc::launch and ihc::collect calls appear as blocking streaming write and streaming read operations.

**Interaction with External Systems**

Task functions can use a global instance of the ihc::stream_in class to take an input from the external system, or a global instance of the ihc::stream_out class to provide output to the external system.

The global ihc::stream_in and ihc::stream_out streams must be declared outside of any struct variables, and they cannot be declared in arrays.
Communication Between HLS Task Functions

For two task functions to communicate with each other, connect them with a global ihc::stream object (instead of the ihc::stream_in and ihc::stream_out objects).

The global ihc::stream object must be declared outside of any struct variables, and it cannot be declared in an array.

The ihc::stream object has an API very similar to the ihc::stream_in and ihc::stream_out classes. However, since these streams always require handshaking, the API does not support the parameters ihc::usesReady or ihc::usesValid. They do support tryRead and tryWrite API functions.

The ihc::stream objects can have both of their endpoints within the system of tasks. This includes within the same function as well. For an example of using an ihc::stream within a single function as a FIFO, see the following tutorial:

<quartus_installdir>/hls/examples/tutorials/system_of_tasks/internal_stream

If an instance of the ihc::stream class has only one endpoint within the system of tasks, it is treated as if it were a ihc::stream_in or ihc::stream_out class based on its usage within the system, so it can be used interchangeably with ihc::stream_in or ihc::stream_out (provided that the limitations do not affect the design). An ihc::stream object can be used for multiple tasks to communicate with one another. See the following tutorial:

<quartus_installdir>/hls/examples/tutorials/system_of_tasks/parallel_loops

The following diagram shows how you might use the ihc::stream object to communicate between task functions:
**Figure 13. Example of a Systems of Tasks using Internal Streams**

```
ihc::stream<> internal;
//global scope

Component Function

component top
{
    ihc::launch

    ihc::launch
    //no collect
}

Task Function

producer
{
    for (...) {
        internal.write()
    }
}

Task Function

consumer
{
    for (...) {
        internal.read()
    }
}
```

**HLS Task Function Restrictions**

HLS task functions are subject to the following restrictions:

- Task functions cannot be shared between multiple components.
- All read sites and write sites for a stream must be within the same function (component or task).
- If you implement a class member function as a task function, the member function must be static. If you want to parameterize the function behavior, use function parameters or template parameters. You cannot use instance variables of an object.
- A task function can be launched (with `ihc::launch`) only from one component function or task function. The launching function and the collecting function can be different functions but they must part of the same component system of tasks.
- A task function can be collected (with `ihc::collect`) only from one component or task function. The collecting function and the launching function can be different functions but they must part of the same component system of tasks.
- No guarantee of execution order is provided between independent I/O instructions, even at the task level.

The `ihc::launch` and `ihc::collect` calls to a particular task function are executed in order.

Any stream accesses to that task from the current function are executed in instruction order only with respect to `ihc::launch` and `ihc::collect` calls to the corresponding function.

**Figure 14. Example 1 of a Valid `ihc::launch/ihc::collect` Sequence**
Figure 15. Example 2 of a Valid `ihc::launch/ihc::collect` Sequence

```
foo {
    ihc::launch
    ihc::launch
    ihc::collect
    ihc::collect
}
```
Figure 16. Example 3 of a Valid `ihc::launch/ihc::collect` Sequence

```
Component Function

component baz
{
    ihc::launch
    ihc::launch
}

Task Function

Task Function

foo
{
    ihc::launch
    ihc::launch
}

Task Function

bar
{
    ihc::collect
    ihc::collect
}
```

10. Systems of Tasks
Figure 17. Example of an Invalid `ihc::launch/ihc::collect` Sequence

```
foo {
  ihc::launch
  Task Mul
  ihc::collect
}

bar {
  ihc::launch
  ihc::collect
}
```

**Task Attributes**

You can use the following function-level attributes on an HLS task function:

- `hls_max_concurrency`
- `hls_component_ii`
- `hls_scheduler_target_fmax_mhz`
- `hls_disable_component_pipelining`

In addition to these function attributes, you can use any HLS attributes and pragmas within your HLS task functions. For example, you can use attributes and pragmas like `#pragma ii`, `#pragma ivdep`, `hls_memory`, and `hls_register`.

You cannot use component macros or component invocation interface control attributes when you define HLS task functions. For example, you cannot use `hls_avalon_agent_register_argument`, `hls_conduit_argument`, `hls_stall_free_return`, or `hls_avalon_streaming_component`.

**Task Capacity**

The compiler schedules launches and collects in the top-level component or caller datapath.

If you schedule multiple tasks with different latencies, you might need to add `capacity` to `ihc::launch` and `ihc::collect` calls.
Similarly, setting the capacity parameter in an ihc::launch call inserts a buffer that allows the caller to push work onto the task function which is then free to pull work off that queue when it can.

Setting the capacity parameter in an ihc::collect call inserts a buffer that can hold the return values as they are computed by the task function. The caller function is free to pull the return values from the buffer at a convenient later time without causing backpressure on the task function.

For more information on how and when to use the capacity parameter with your ihc::launch and ihc::collect calls, refer to "Explicitly Add Buffer Capacity to Your Design When Needed" in the Intel HLS Compiler Best Practices Guide.

You can also review the following tutorial:

<quartus_installdir>/hls/examples/tutorials/system_of_tasks/launch_and_collect_capacity

Related Information

- Avalon Streaming Interfaces on page 25
- Avalon Memory-Mapped Host Interfaces on page 32
- Pipes on page 28

10.2. Internal Streams

You can use the HLS ihc::stream object as a FIFO in a single task or component.

For an example of using the HLS tasks ihc::stream object as a FIFO, review the tutorial in <quartus_installdir>/hls/examples/tutorials/system_of_tasks/internal_stream.

To help you understand the tutorial better, review the following diagram showing a store-load dependency:

This diagram is simplified from the tutorial. It shows 10 iterations, while the tutorial goes through 32 iterations.

In the diagram, i is the index of the outer loop and j is the index of the inner loop.

Each iteration of the outer loop reads all the values written by the previous loop iteration and writes one less value to the buffer. The internal stream outperforms the array in this design because array must allocate enough space to store written values before the values are read, but an internal stream does not need to allocate this space.
In addition, the trip count of the inner loop decreases by one in each outer loop, so the space claimed by array is never filled after the first iteration, which wastes area.

10.3. System of Tasks Simulation

When you simulate a system of tasks design where the completion of a task function is not synchronized with an \texttt{ihc::collect} call, use the \texttt{ihc\_hls\_set\_component\_wait\_cycle} testbench API function to allow output from that task function to be returned after the component function finishes running.

By default, the simulation process simulates an additional 100 cycles after a component asserts the \texttt{done} signal to ensure all operations have propagated back to the testbench. This function tells the simulation process for the specified component to continue running for the specified number of cycles in addition to the default wait period of 100 cycles.

If you do not use this function in your testbench, the latency of some task functions might make your simulation output inaccurate or cause your simulation testbench to hang.

For an example of a valid systems of task design where the completion of a task function is not synchronized with an \texttt{ihc::collect} call, see Example 3 of a Valid \texttt{ihc::launch/ihc::collect} Sequence.

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>\texttt{ihc_hls_set_component_wait_cycle}</td>
<td>This function tells the simulation process to continue running for a specified number of additional cycles (beyond the default wait period of 100 cycles) after the \texttt{done} signal for the specified component is observed.</td>
</tr>
</tbody>
</table>
11. Libraries

With libraries, you can reuse functions without knowing the underlying hardware design or implementation details. Libraries can be created with Intel FPGA high-level design tools including the Intel HLS Compiler and the Intel FPGA SDK for OpenCL*, either from code initially targeting that tool or from RTL code.

The Intel HLS Compiler supports two types of libraries:

- **Static-object libraries**
- **Source code libraries**

**Static-Object Libraries**

A static-object library is a single platform-specific archive file that contains one or more object files. A static-object file contains implementations of one or more functions. The object and library files use the same formats as the operating system that you compile your Intel HLS Compiler code on, with additional sections that carry HLS-specific information.

On Linux platforms, an object library is a `.a` archive file that contains `.o` object files. On Windows platforms, a library is a `.lib` archive file that contains `.obj` object files.

A static-object library includes one or more function signature files that you include in your component source code so that your component can call the functions provided by the library. A function signature file is a C-style header file (`.h`) that declares the signatures of the functions that are provided in an object library.

Static-object libraries can be created from RTL or high-level source code.

**Source Code Libraries**

A source code library is a C-style header file that contains a source code library. You include this header file in your component source code, and the header file code is compiled along with your component.

You can use C++ templates to make your source code library more customizable.

The Intel HLS Compiler provides some source code libraries that provide you with FPGA-optimized code for some commonly-used algorithms.

For details about source code libraries included with the Intel HLS Compiler, see the following sections:

- **Arbitrary Precision Math Support** on page 79
- **Advanced Math Source Code Libraries** on page 198

*Other names and brands may be claimed as the property of others.
11.1. Static-Object Libraries

A static-object library is a single platform-specific archive file that contains one or more object files, each of which contains implementations of one or more functions.

The object and library files use the same formats as the operating system that you compile your Intel HLS Compiler code on, with additional sections that carry additional library information. On Linux platforms, a library is a `.a` archive file that contains `.o` object files. On Windows platforms, a library is a `.lib` archive file that contains `.obj` object files.

You can call the functions in the library from your component without needing to know the hardware design or the implementation details of the underlying functions in the library. Add the library to the `i++` command line when you compile your component.

You can create a library from your HLS C++ code source files or register transfer level (RTL) language source files. You can target the library for use with one of the following Intel high-level design products:

- Intel HLS Compiler Pro Edition
- Intel FPGA SDK for OpenCL Pro Edition

To create a library from your HLS code that targets the Intel FPGA SDK for OpenCL, you must have the Intel FPGA SDK for OpenCL Pro Edition installed. The version of the SDK must be the same as your version of Intel HLS Compiler.

Creating a library is a two-step process

1. Each object file is generated from input source files with the `fpga_crossgen` command.
   
   The required input source files depend on the type of source code you are creating the object from.
   
   An object is effectively an intermediate representation of your source code with both a CPU representation and an FPGA representation of your code.
   
   An object can be targeted for use with only one Intel high-level design product. If you want to target more than one high-level design product, you must generate a separate object for each target product.

2. Object files are collected into a library file with the `fpga_libtool` command
   
   Objects created from different types of source code can be collected into a library, provided all objects target the same high-level design product.
   
   Libraries must be built and used by the same version number Intel FPGA high-level design tool. For example, to compile your component with the Intel HLS Compiler Version 21.2, the libraries included in your component must have been created with a version 21.2 Intel FPGA high-level design tool.
11.2. Creating a Static-Object Library

Creating a static-object library is a multistep process where you create the objects that you want to include in a library and then collect the objects into a library file.

To create a static-object library:

1. Create the objects for your library with the `fpga_crossgen` command. You can create your objects from a variety of sources:
   - Create an object from HLS code.
     For details, see Creating Objects From HLS Code on page 114.
   - Create an object from RTL code.
     For details, see Creating Objects From RTL Code on page 117.
   - Create an object from OpenCL code.
     For details, see Creating Library Objects From OpenCL Code in the Intel FPGA SDK for OpenCL Pro Edition Programming Guide.

2. Collect the objects into an object library with the `fpga_libtool` command.
   For details, see Packaging Object Files Into a Library on page 137.

For example, if you wanted to create a Linux HLS object library called `foobar` from HLS code in a file called `foo.cpp` and OpenCL code in a file called `bar.cl`, run the following commands:

```bash
fpga_crossgen foo.cpp -target hls -o foo.o
fpga_crossgen bar.cl -target hls -o bar.o
fpga_libtool -target hls -create foobar.a foo.o bar.o
```

You can use the resulting library (`foobar.a`) in your component by including the header file or files that you created for the library (for example, `foobar.a`, or `foo.h` and `bar.h`) in your component.
When you compile your component, specify the library in the option of the `i++` command. For example, to compile a component `baz.cpp` that uses the `foobar.a` library, issue the following command:

```
i++ baz.cpp foobar.a
```

### 11.3. Creating Objects From HLS Code

You can create a library from object files from your HLS source code. An HLS-based object file contains code for CPU execution (testbench and emulation) and FPGA execution. A library can contain multiple object files.

You can create object files for use with different Intel high-level design tools from the same HLS source code.

Depending on the target high-level design tool, your source code might require adjustments to support tool-specific data types or constructs.

**Remember:** Ensure that the library functions in your HLS source code are not declared with the `HLS component` keyword.

**Intel HLS Compiler**

No additional work is needed in your HLS source code when you use the code to create objects for Intel HLS Compiler libraries.

**Intel FPGA SDK for OpenCL**

The Intel FPGA SDK for OpenCL supports language constructs that are not natively supported by C++. Your code might need modifications to support those constructs.

The Intel HLS Compiler supports a limited set of OpenCL language constructs through the `ocl_types.h` header file. For details, review [Supported OpenCL Language Constructs](#) on page 116.

To create an object from your HLS code that targets the Intel FPGA SDK for OpenCL, you must have the Intel FPGA SDK for OpenCL Pro Edition installed. The version of the SDK must be the same as your version of Intel HLS Compiler.

### 11.3.1. Creating an Object File From HLS Code

Use the `fpga_crossgen` command to create objects for your library from your HLS code. An object created from HLS code contains information required both for emulating the functions in the object and synthesizing the hardware for the object functions.

**Restriction:** Creating object files from HLS code is supported only on Linux operating systems.

The `fpga_crossgen` command creates one object file from one input source file. The object created can be used only libraries that target the same Intel high-level design tool.
Objects are assigned the same version number as the version number of your Intel HLS Compiler installation. Libraries can contain only objects with the same version number, and can only be used with Intel high-level design tools with the same version number.

All functions in your HLS code that need to be exported to a library must be defined using the `HLS_EXTERNAL` macro. For example:

```c
extern "C" HLS_EXTERNAL int my_hls_func(int x);
```

**Remember:** Ensure that the functions in your HLS source code are not declared with the HLS component keyword.

1. Create a library object with the following command:

   ```bash
   fpga_crossgen <source_file> --target target_HLD_tool [-o <object_file_name>]
   ```

   Where the command parameters are defined as follows:
   
   - `target_HLD_tool`
     
     The target Intel high-level design tool for this library. This parameter can have one of the following values:
     
     - `hls`
       
       Target this object to be included in libraries for components developed with the Intel HLS Compiler.
       
       Objects built for the Intel HLS Compiler are created as operating system specific object files (.o on Linux). You cannot use objects created on one operating system with the Intel HLS Compiler running on a different operating system.
     
     - `aoc`
       
       Target this object to be included in libraries for kernels developed with the Intel FPGA SDK for OpenCL.
       
       Objects built for the Intel FPGA SDK for OpenCL are not operating system specific. The objects are created as Intel FPGA SDK for OpenCL object files (.aoco).
       
       You must have the Intel FPGA SDK for OpenCL Pro Edition installed to use this option. The version of the SDK must be the same as your version of Intel HLS Compiler.
       
   If you do not specify an object file name with the `-o` option, the object file name defaults to be the same name as the source file name.

   After you generated all objects that you want to include in your library, collect the objects in the library with the `fpga_libtool` command. For details, see Packaging Object Files Into a Library on page 137.
11.3.2. Supported OpenCL Language Constructs

If you are using the Intel HLS Compiler to develop libraries to use with the Intel FPGA SDK for OpenCL, you might need access to OpenCL language constructs that are not typically available natively from C++ language elements. The Intel HLS Compiler provides support for some OpenCL language constructs through the ocl_types.h header file.

All basic signed and unsigned OpenCL data types (double, float, long long, long, int, short, char and bool) are supported without needing the ocl_types.h header file.

Add OpenCL language construct support by adding the following code to your component:

```c
#include "HLS/ocl_types.h"
```

The OpenCL types header file adds support for the following OpenCL language constructs:

- OpenCL address space qualifiers
- Arbitrary precision integers (up to 64 bits)
- OpenCL vector data types

**Note:**

The size of the long data type is 64 bits in OpenCL and for the Intel HLS Compiler on Linux systems. For Intel HLS Compiler on Windows systems, the size of long is 32 bits.

### OpenCL Address Space Qualifiers

The ocl_types.h header file adds macros to support defining pointer in different OpenCL address spaces as follows:

<table>
<thead>
<tr>
<th>OpenCL Address Space Qualifier</th>
<th>Intel HLS Compiler Macro</th>
</tr>
</thead>
<tbody>
<tr>
<td>__global</td>
<td>OCL_ADDRSP_GLOBAL</td>
</tr>
<tr>
<td>__local</td>
<td>OCL_ADDRSP_LOCAL</td>
</tr>
<tr>
<td>__constant</td>
<td>OCL_ADDRSP_CONSTANT</td>
</tr>
<tr>
<td>__private</td>
<td>OCL_ADDRSP_PRIVATE</td>
</tr>
</tbody>
</table>

### Arbitrary Precision Integers

The ocl_types.h header file supports the OpenCL intX_t and uintX_t data types up to 64 bits. However, these data types are in the ihc namespace to avoid conflicts with C-system header definitions.

That is, you can use ihc::int1_t through to ihc::int64_t and ihc::uint1_t through to ihc::uint64_t in your component.

Only use these data types to exchange data on your component interface (for example, parameters). Assign them to HLS ac_int<> data types in your component code.
Restriction: While you can use a `using namespace ihc;` declaration in your code, power-of-two arbitrary precision datatypes still require the `ihc::` namespace prefix. For example, you must declare `ihc::uint8_t` whether it preceded by a `using namespace ihc;` declaration or not.

11.4. Creating Objects From RTL Code

You can create a library from object files that package register transfer level (RTL) language source files. An RTL-based object file also contains an object manifest file (in XML format) that identifies the functions that are callable in the object file. A library can contain multiple RTL-based objects.

Creating a library from RTL code is a two-step process. First, each object file is created from the RTL source and emulation models as described in the object manifest file with the `fpga_crossgen` command. Then, one or more object files are collected into an HLS library file with the `fpga_libtool` command.

![High-Level View of the Library Creation Process from RTL](image)

To create a library from RTL code, you need to create the following files and components:

Table 28. Files and Components Required for Creating a Library From RTL Code

<table>
<thead>
<tr>
<th>File or Component</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RTL-based Functions</strong></td>
<td></td>
</tr>
<tr>
<td>RTL module source files</td>
<td>Verilog (.v), System Verilog (.sv), or VHDL (.vhd) files and accompanying</td>
</tr>
<tr>
<td></td>
<td>memory initialization files (.mif or .hex) that define the RTL modules in</td>
</tr>
<tr>
<td></td>
<td>the library. You cannot use additional files such as Intel Quartus Prime</td>
</tr>
<tr>
<td></td>
<td>IP File (.qip), Synopsys Design Constraints File (.sdc), or Tcl Script File</td>
</tr>
<tr>
<td></td>
<td>(.tcl).</td>
</tr>
</tbody>
</table>

continued...
### 11.4.1. RTL Modules and the HLS Pipeline

HLS libraries allows you to use RTL modules that are written in Verilog, SystemVerilog, or VHDL inside HLS components. The Intel HLS Compiler integrates the RTL modules into the HLS pipeline architecture.

Consider using HLS libraries in the following situations:
- You want to use optimized and verified RTL modules in HLS components without rewriting the modules as C++ functions.
- You want to implement HLS component functionality that you cannot express effectively in C++.

#### 11.4.1.1. Integration of an RTL Module into the HLS Pipeline

When you specify an HLS library during component compilation, the Intel HLS Compiler integrates the RTL module within the library into the overall component pipeline.

The following figure shows how an HLS library called myMod might be integrated into the example pipeline described in Intel HLS Compiler Hardware Model on page 14.
Figure 20. Example of Pipeline Architecture That Integrates an HLS Library

```c
extern "C" int myMod (int) ;
component int pe
   (int A, int B, int C) {
   int product1 = A * B;
   int product2 = B * C;
   int mod_output = myMod(C);
   int sum = product1 + product2;
   int result = sum + mod_output;
   return result;
}
```

The depicted RTL module has a latency of 3 cycles. Since the multiply and add operations have a latency of just one cycle, the compiler inserts buffering to balance the latency of the parallel data paths in the pipeline. A balanced latency allows the invocations of the HLS component to execute without stalling the pipeline.

Specifying the latency of the RTL module in the HLS library object manifest file allows the HLS compiler to balance the pipeline latencies in the HLS component. The pipeline integration protocol uses ready/valid handshaking, so the latency of the RTL module can be variable. However, the variability in the latency should be small to maximize performance. In addition, specify the latency in the HLS library object manifest file for the object in the HLS library so that the RTL module experiences a good approximation of the actual latency in steady state.

**Note:** You must specify the RTL module latency correctly in the HLS library object manifest file, or you get bad quality of results (QoR) for your component.

### 11.4.1.2. RTL Module Interfaces

For an RTL module to properly interact with other compiler-generated operations, you must support a simple ready/valid handshaking protocol at both the input and the output of an RTL module.

An RTL module must use a single streaming interface. That is, a single pair of ready and valid logic must control all the inputs.

You have the option to provide the necessary streaming ports but declare the RTL module as stall-free. In this case, you do not have to implement proper stall behavior because the Intel HLS Compiler creates a wrapper for your module.
You must handle invalid signals properly if your RTL module has an internal state. For more information, see Stall-Free RTL on page 134.

Consider the following interfaces for the RTL module myMod:

In this diagram, myMod interacts with the upstream module through data signals, arg1 and arg2, and control signals, invalid (input) and oready (output). The invalid control signal equals 1 (invalid = 1) if and only if data signal arg1 and data signal arg2 contain valid data. When the control signal oready equals 1 (oready = 1), it indicates that the myMod RTL module can process the data signals arg1 and arg2 if they are valid (that is, invalid = 1). When invalid = 1 and oready = 0, the upstream module holds the values of invalid, arg1, and arg2 in the next clock cycle.

The myMod module interacts with the downstream pipeline logic through the data signal result and the control signals, ovalid (output) and iready (input). The ovalid control signal equals 1 (ovalid = 1) if and only if the data signal result contains valid data. When the iready control signal equals 1 (iready = 1), the downstream module can process the data signal result if it is valid. When ovalid = 1 and iready = 0, the myMod RTL module must hold the valid of the ovalid and result signals in the next clock cycle.

11.4.1.2.1. RTL Module Interface Signals

The Intel HLS Compiler expects the RTL module to support a single interface with readyLatency = 0, at both input and output.

As shown in RTL Module Interfaces on page 119, the RTL module must have four ports:
- invalid and iready as the input ready/valid interface
- ovalid and oready as the output ready/valid interface

The following figure illustrates the timing diagram for input data transfer with back pressure.
For an RTL module with a fixed latency, the output signals (\texttt{valid} and \texttt{ready}) can have constant high values, and the input ready signal (\texttt{iready}) can be ignored.

A stall-free RTL module might receive an invalid input signal (\texttt{invalid} is low). In this case, the module ignores the input and produces invalid data on the output. For a stall-free RTL module without an internal state, it might be easier to propagate the invalid input through the module. However, for an RTL module with an internal state, you must handle an \texttt{invalid} = 0 input carefully.
Example Timing Diagram of a Stall-free RTL Component

Consider the following example timing diagram of a stall-free RTL component:

Figure 22. Timing Diagram of a Stall-free RTL Component

For this component, the following ATTRIBUTE elements are set in the object manifest file for the RTL module:

- **IS_STALL_FREE value = "yes"
- **IS_FIXED_LATENCY value = "yes"
- **EXPECTED_LATENCY value = "2"
Example Timing Diagram of a Non-stall-free RTL Component

Consider the following example timing diagram of a non-stall-free RTL component:

**Figure 23. Timing Diagram of a Non-stall-free RTL Component**

For this component, the following ATTRIBUTE elements are set in the object manifest file for the RTL module:

- **IS_STALL_FREE** value = "no"
- **IS_FIXED_LATENCY** value = "no"
- **EXPECTED_LATENCY** value = "4"

11.4.1.3. RTL Reset and Clock Signals

Resets and clocks of RTL modules are connected to the same clock and reset drivers as the rest of the HLS pipeline.

Because of the common clock and reset drivers, an RTL module runs in the same clock domain as the HLS component that is integrating the RTL module. The module reset input is asserted whenever the HLS component is reset.

11.4.1.3.1. Intel Agilex and Intel Stratix 10 Design-Specific Reset Requirements for Stall-Free and Stallable RTL Modules

When you create an RTL module for Intel Agilex and Intel Stratix 10 HLS designs, ensure that the module satisfies specific logic reset requirements.
Reset Requirements for Stall-Free RTL Modules

A stall-free RTL module is a fixed-latency module for which the Intel HLS Compiler can optimize away stall logic.

- When creating a stall-free RTL module for an Intel Agilex and Intel Stratix 10 design, use synchronous clear signals only.
- After deassertion of the reset signal to the stall-free RTL module, the module must be operational within 15 clock cycles. If the reset signal is pipelined within the module, this requirement limits the reset pipelining to no more than 15 stages.

Reset Requirements for Stallable RTL Modules

A stallable RTL module has a variable latency, and it relies on backpressured input and output interfaces to function correctly.

- When creating a stallable RTL module for an Intel Agilex and Intel Stratix 10 design, use synchronous clear signals only.
- After assertion of the reset signal to the stallable RTL module, the module must deassert its oready and ovalid interface signals within 40 clock cycles.
- After deassertion of the reset signal to the stallable RTL module, the module must be fully operational within 40 clock cycles. The module signals its readiness by asserting the oready interface signal.

11.4.1.4. Object Manifest File Syntax

The HLS library object manifest file is an XML file that maps the RTL modules in a library object to functions that can be called by your HLS code. The Intel HLS Compiler uses the properties defined in the manifest file to integrate an RTL module into the component pipeline.

The following example shows a simple object manifest file for an RTL module that implements a double-precision square root function. The RTL module is implemented in VHDL with a Verilog wrapper.

The following object manifest file is for an RTL module named `my_fp_sqrt_double` (line 2) that implements a callable function with a C interface named `my_sqrtfd` (line 2).

```xml
1: <RTL_SPEC>
2:   <FUNCTION name="my_sqrtfd" module="my_fp_sqrt_double">
3:     <ATTRIBUTES>
4:       <IS_STALL_FREE value="yes"/>
5:       <IS_FIXED_LATENCY value="yes"/>
6:       <EXPECTED_LATENCY value="31"/>
7:       <CAPACITY value="31"/>
8:       <HAS_SIDE_EFFECTS value="no"/>
9:       <ALLOW_MERGING value="no"/>
10:      <PARAMETER name="WIDTH" value="64"/>
11:     </ATTRIBUTES>
12:     <INTERFACE>
13:       <AVALON port="clock" type="clock"/>
14:       <AVALON port="resetn" type="resetn"/>
15:       <AVALON port="ivalid" type="ivalid"/>
16:       <AVALON port="iready" type="iready"/>
17:       <AVALON port="ovalid" type="ovalid"/>
18:       <AVALON port="oready" type="oready"/>
19:       <INPUT port="datain" width="64"/>
20:       <OUTPUT port="dataout" width="64"/>
```

MNL-1083 | 2021.06.21
124
Table 29. Elements and Attributes in the Object Manifest File

<table>
<thead>
<tr>
<th>XML Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RTL_SPEC</td>
<td>Top-level element in the object manifest file. There can only be one such top-level element in the file.</td>
</tr>
<tr>
<td>FUNCTION</td>
<td>Element that defines the HLS function that the RTL module implements. The name attribute within the FUNCTION element specifies the function name. You might have multiple FUNCTION elements, each declaring a different function that you can call from the HLS component. The same RTL module can implement multiple functions by specifying different parameters. To use the same module with different parameter combinations, create a separate FUNCTION tag for each parameter combination.</td>
</tr>
<tr>
<td>ATTRIBUTES</td>
<td>Element that contains other XML elements that describe various characteristics (for example, latency) of the RTL module. The example RTL module takes one PARAMETER setting named WIDTH, which has a value of 64. See Table 30 on page 127 for more details other ATTRIBUTES-specific elements. If you create multiple RTL-based functions using different modules or use the same RTL module with different PARAMETER settings, you must create a separate FUNCTION element for each function.</td>
</tr>
</tbody>
</table>
### XML Element Description

<table>
<thead>
<tr>
<th>XML Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTERFACE</td>
<td>Element that contains other XML elements that describe the RTL module interface. The example object manifest file shows the streaming interface signals that every RTL module must provide (that is, clock, resetn, invalid, iready, ovalid, and oready). The resetn signal is active low. Its synchronicity depends on the target device:</td>
</tr>
<tr>
<td></td>
<td><strong>Intel Agilex</strong> The resetn signal is synchronous to the clock signal. For more information about reset signal timing, see Intel Agilex and Intel Stratix 10 Design-Specific Reset Requirements for Stall-Free and Stallable RTL Modules on page 123.</td>
</tr>
<tr>
<td></td>
<td><strong>Intel Arria® 10</strong> The resetn signal is asynchronous to the clock signal.</td>
</tr>
<tr>
<td></td>
<td><strong>Intel Stratix 10</strong> The resetn signal is synchronous to the clock signal. For more information about reset signal timing, see Intel Agilex and Intel Stratix 10 Design-Specific Reset Requirements for Stall-Free and Stallable RTL Modules on page 123.</td>
</tr>
<tr>
<td></td>
<td>The signal names must match the ones specified in the RTL module properties file. An error occurs during library creation if a signal name is different in the RTL code and the RTL module properties file.</td>
</tr>
<tr>
<td>REQUIREMENTS</td>
<td>Element that specifies one or more RTL resource files (that is, .v, .sv, .vhd, .hex, and .mif). The specified paths to these files are relative to the location of the object manifest file. Each RTL resource file becomes part of the associated Platform Designer component that corresponds to the entire HLS component. HLS libraries do not support .qip files.</td>
</tr>
<tr>
<td>RESOURCES</td>
<td>Optional element that specifies an estimate of the FPGA resources that the RTL module uses. If you do not specify this element, the estimated FPGA resources that the RTL module uses defaults to zero in the HLS resource estimation report.</td>
</tr>
</tbody>
</table>

### 11.4.1.4.1. XML Elements for ATTRIBUTES

In the RTL module properties file of the RTL module within an HLS library, there are XML elements under ATTRIBUTES that you can specify to set module characteristics.
Table 30. XML Elements for the RTL module properties file ATTRIBUTES Element

<table>
<thead>
<tr>
<th>XML Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS_STALL_FREE</td>
<td>Instructs the Intel HLS Compiler to remove all stall logic around the RTL module. Set IS_STALL_FREE to &quot;yes&quot; to indicate that the module does not generate stalls internally and it cannot properly handle incoming stalls. The module ignores the stall input. If you set IS_STALL_FREE to &quot;no&quot;, the module must properly handle all stall and valid signals. If you set IS_STALL_FREE to &quot;yes&quot;, you must also set IS_FIXED_LATENCY to &quot;yes&quot;. Also, if the RTL module has an internal state, it must properly handle invalid=0 inputs.</td>
</tr>
<tr>
<td>IS_FIXED_LATENCY</td>
<td>Indicates whether the RTL module has a fixed latency. Set IS_FIXED_LATENCY to &quot;yes&quot; if the RTL module always takes a known number of clock cycles to compute its output. The value you assign to the EXPECTED_LATENCY element specifies the number of clock cycles. The safe value for IS_FIXED_LATENCY is &quot;no&quot;. When you set IS_FIXED_LATENCY=&quot;no&quot;, the EXPECTED_LATENCY value must be at least 1. For a given RTL module, you may set IS_FIXED_LATENCY to &quot;yes&quot; and IS_STALL_FREE to &quot;no&quot;. Such a module produces its output in a fixed number of clock cycles and handles stall signals properly.</td>
</tr>
<tr>
<td>EXPECTED_LATENCY</td>
<td>Specifies the expected latency of the RTL module. If you set IS_FIXED_LATENCY to &quot;yes&quot;, set the EXPECTED_LATENCY value to be the exact latency of the module. Otherwise, the Intel HLS Compiler generates incorrect hardware. For a module with variable latency, the Intel HLS Compiler balances the pipeline around this module to the EXPECTED_LATENCY value that you specify. For modules that can stall and require use of signals such as iready, the EXPECTED_LATENCY value must be set to at least 1. The specified value and the actual latency might differ for a module with variable latency, which might affect the number of stalls inside the pipeline. However, the resulting hardware is functionally correct.</td>
</tr>
<tr>
<td>CAPACITY</td>
<td>Specifies the number of multiple inputs that this module can process simultaneously. You must specify a value for CAPACITY if you also set IS_STALL_FREE=&quot;no&quot; and IS_FIXED_LATENCY=&quot;no&quot;. Otherwise, you do not need to specify a value for CAPACITY. If CAPACITY is strictly less than EXPECTED_LATENCY, the Intel HLS Compiler automatically inserts capacity-balancing FIFO buffers after this module when necessary. A conservative but safe value for CAPACITY is 1.</td>
</tr>
<tr>
<td>XML Element</td>
<td>Description</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>HAS_SIDE_EFFECTS</td>
<td>Indicates whether the RTL module has side effects. Modules that have internal states or communicate with external memories are examples of modules with side effects. Set HAS_SIDE_EFFECTS to &quot;yes&quot; to indicate that the module has side effects. Specifying HAS_SIDE_EFFECTS to &quot;yes&quot; ensures that optimization efforts do not remove calls to modules with side effects. Stall-free modules with side effects (that is, IS_STALL_FREE=&quot;yes&quot; and HAS_SIDE_EFFECTS=&quot;yes&quot;) must properly handle invalid=0 input cases because the module might receive invalid data occasionally. A conservative but safe value for HAS_SIDE_EFFECTS is &quot;yes&quot;. This element along with the ALLOW_MERGING element allow the Intel HLS Compiler to perform certain optimizations. For details, see Interaction Between ALLOW_MERGING and HAS_SIDE_EFFECTS Elements on page 128.</td>
</tr>
<tr>
<td>ALLOW_MERGING</td>
<td>Indicates that the compiler can merge multiple instances of this RTL module. Set ALLOW_MERGING to &quot;yes&quot; to allow merging of multiple instances of the module. Intel recommends setting ALLOW_MERGING to &quot;yes&quot;. The safe value for ALLOW_MERGING is &quot;no&quot;. Marking the module with HAS_SIDE_EFFECTS=&quot;yes&quot; does not prevent merging. This element along with the HAS_SIDE_EFFECTS element allow the Intel HLS Compiler to perform certain optimizations. For details, see Interaction Between ALLOW_MERGING and HAS_SIDE_EFFECTS Elements on page 128.</td>
</tr>
<tr>
<td>PARAMETER</td>
<td>Specifies the value of an RTL module parameter. PARAMETER attributes: • name Specifies the name of the RTL module parameter. • value Specifies a decimal numeric value for the parameter. The value for an RTL module parameter can be specified using either a value or a type attribute.</td>
</tr>
</tbody>
</table>

**Interaction Between ALLOW_MERGING and HAS_SIDE_EFFECTS Elements**

The ALLOW_MERGING and HAS_SIDE_EFFECTS elements help the Intel HLS Compiler optimize hardware use of your RTL component. Consider the effects of these element carefully when setting their values.

The combination you select depends on your design architecture. For example, you can tell the compiler that you want the compiler to replicate the RTL block for multiple calls or vectorized code.
### Table 31. ALLOW_MERGING and HAS_SIDE_EFFECTS Combinations and Their Effect

<table>
<thead>
<tr>
<th>ALLOW_MERGING Value</th>
<th>HAS_SIDE_EFFECTS Value</th>
<th>Combination Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>no</td>
<td>no</td>
<td>Each call to an RTL library corresponds to one distinct instance in the hardware. Calls might be optimized away by the compiler if deemed redundant or unnecessary. Calls might be vectorized, with multiple instances in the hardware created for a single RTL library call.</td>
</tr>
<tr>
<td>no</td>
<td>yes</td>
<td>Each call to an RTL library corresponds to one distinct instance in hardware. Calls are not optimized away by the compiler.</td>
</tr>
<tr>
<td>yes</td>
<td>no</td>
<td>Multiple calls to an RTL library might be merged into one call, and hence correspond to one instance in the hardware. Calls might be optimized away by the compiler if deemed redundant or unnecessary. Calls might be vectorized, with multiple instances in the hardware created for a single RTL library call.</td>
</tr>
<tr>
<td>yes</td>
<td>yes</td>
<td>Multiple calls to an RTL library might be merged into one call, and hence correspond to one instance in hardware. Calls are not optimized away by the compiler.</td>
</tr>
</tbody>
</table>

### 11.4.1.4.2. XML Elements for INTERFACE

In the RTL module properties file of the RTL module within an HLS library, there are XML elements under INTERFACE that define aspects of the RTL module interface.

The RTL module cannot access the memories of the HLS component.

### Table 32. Mandatory XML Elements for the RTL module properties file INTERFACE Element

<table>
<thead>
<tr>
<th>XML Element</th>
<th>Description</th>
</tr>
</thead>
</table>
| INPUT       | Specifies the input parameter of the RTL module that receives the value of a call argument with the RTL-based function is called. INPUT attributes:  
  * port  
    Specifies the port name of the RTL module.  
  * width  
    Specifies the width of the port in bits.  
  All call arguments must be passed by value. You cannot use reference, pointer, and array type arguments. |
| OUTPUT      | Specifies the output parameter of the RTL module that represents the return value of functions based on this module. OUTPUT attributes:  
  * port  
    Specifies the port name of the RTL module.  
  * width  
    Specifies the width of the port in bits.  
    The width of the port must match the size of the C datatype of the function return value. The port width is always a multiple of 8 bits.  
    The return value cannot be a pointer. |
<p>| STREAM      | Specifies the stream parameters to the RTL module. STREAM attributes: |</p>
<table>
<thead>
<tr>
<th>XML Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>• port</td>
<td>Specifies the base port name for the stream input to the RTL module.</td>
</tr>
<tr>
<td>• width</td>
<td>Specifies the width of the port in bits.</td>
</tr>
<tr>
<td>• direction</td>
<td>Specifies the streaming direction. Valid values for the direction attribute are in and out. Use direction=&quot;in&quot; for streams that are read by the RTL module and direction=&quot;out&quot; for streams that written to by the RTL module.</td>
</tr>
<tr>
<td>• buffer</td>
<td>Specifies the size of the stream buffer in words. The default value is buffer=&quot;0&quot;.</td>
</tr>
<tr>
<td>• usesValid</td>
<td>Specifies whether this RTL module uses the valid signal for the stream interface. The default value is usesValid=&quot;no&quot;.</td>
</tr>
<tr>
<td>• usesReady</td>
<td>Specifies whether this RTL module uses the ready signal for the stream interface. The default value is usesReady=&quot;no&quot;.</td>
</tr>
<tr>
<td>• usesEmpty</td>
<td>Specifies whether this RTL module uses the empty signal for the stream interface. The default value is usesEmpty=&quot;no&quot;.</td>
</tr>
<tr>
<td>• usesPackets</td>
<td>Specifies whether this RTL module uses the packet signals for the stream interface. The default value is usesPackets=&quot;no&quot;.</td>
</tr>
<tr>
<td>• firstSymbolInHighOrderBits</td>
<td>Specifies whether the data symbols in the stream are in big endian order. The default value is firstSymbolInHighOrderBits=&quot;no&quot;.</td>
</tr>
<tr>
<td>• bitsPerSymbol</td>
<td>Specifies how the data is broken into symbols on the data bus. The default value is bitsPerSymbol=&quot;0&quot;.</td>
</tr>
<tr>
<td>• readyLatency</td>
<td>Specifies the number of cycles between when the ready signal is deasserted and when the input stream can no longer accept new inputs. The default value is readyLatency=&quot;0&quot;. The values you specify here must match the values for the stream object input interface parameters in your component. For details about stream input interface parameters in your component, see Intel HLS Compiler Pro Edition Streaming Input Interfaces on page 181. The signal names in your RTL and your component code must align. For details, see Mapping HLS Data Types to RTL Signals on page 131.</td>
</tr>
</tbody>
</table>

11.4.1.4.3. XML Elements for RESOURCES

In the RTL module properties file of the RTL module within an HLS library, there are optional elements under RESOURCES that you can define to specify the estimated FPGA resource utilization of the module. If you do not specify a particular element, it is assigned a default value of zero in the report estimates.
Table 33. XML Elements for the RTL module properties file RESOURCES Element

<table>
<thead>
<tr>
<th>XML Element</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALUTS</td>
<td>Specifies the number of combinational adaptive look-up tables (ALUTs) that the module uses.</td>
</tr>
<tr>
<td>FFS</td>
<td>Specifies the number of dedicated logic registers that the module uses.</td>
</tr>
<tr>
<td>RAMS</td>
<td>Specifies the number of block RAMs that the module uses.</td>
</tr>
<tr>
<td>DSPS</td>
<td>Specifies the number of digital signal processing (DSP) blocks that the module uses.</td>
</tr>
<tr>
<td>MLABS</td>
<td>Specifies the number of memory logic arrays (MLABs) that the module uses. This value is equal to the number of adaptive logic modules (ALMs) that is used for memory divided by 10 because each MLAB consumes 10 ALMs.</td>
</tr>
</tbody>
</table>

11.4.1.5. Mapping HLS Data Types to RTL Signals

All supported composite data types are represented by wide input or output signals. Typically, the components of a composite data type are presented with the first-declared value or value of lowest index in the low-order bits of the signal.

Streams

Stream objects are passed by value to the RTL function. The attributes specified for the STREAM element in the Object Manifest File XML should match the template arguments in the component code.

Your RTL code must use signal names based on the port attribute defined for the STREAM element as follows:

- `<port>_data`
  Signal for the data passed on the stream. This signal must be an input signal for `direction="in"` and an output signal for `direction="out"`.
  The width of this signal must match the `width` attribute in the STREAM element.

- `<port>_valid`
  Used only if `usesValid="yes"` is set in the STREAM element in the XML file. This signal is a single-bit signal for the `valid` signal.
  If `direction="in"`, this signal is an input signal. If `direction="out"`, this is an output signal.

- `<port>_ready`
  Used only if `usesReady="yes"` is set in the STREAM element in the XML file. This signal is a single-bit signal for the `ready` signal.
  If `direction="in"`, this signal is an input signal. If `direction="out"`, this is an output signal.
• `<port>_empty`
  Used only if `usesEmpty="yes"` is set in the STREAM element in the XML file.
  This signal is a single-bit signal for the `empty` signal.
  If `direction="in"`, this signal is an input signal. If `direction="out"`, this is an output signal.

• `<port>_startofpacket`
  Used only if `usesPackets="yes"` is set in the STREAM element in the XML file.
  This signal is a single-bit signal to indicated the start of a packet.
  If `direction="in"`, this signal is an input signal. If `direction="out"`, this is an output signal.

• `<port>_endofpacket`
  Used only if `usesPackets="yes"` is set in the STREAM element in the XML file.
  This signal is a single-bit signal to indicated the end of a packet.
  If `direction="in"`, this signal is an input signal. If `direction="out"`, this is an output signal.

These signal names align with stream object input interface parameter names. For details about stream input interface parameters in your component, see Intel HLS Compiler Pro Edition Streaming Input Interfaces on page 181.

**Arrays**

In C++, arrays are passed as a pointer to the memory in which the array is stored.

The Intel HLS Compiler does not support pointer parameters for RTL modules. However, C++ allows you to pass a struct by value, so you can declare a struct data type that has an array as one of its members and declare your function to accept an argument of this struct-type by value.

**Structs**

You can use both packed and unpacked structs as call arguments and return values in your HLS components and tasks. The members of a struct are presented as slices of the input signal, with the first-declared struct member in the lowest-order bits of the input signal.

• **Unpacked Structs**
  When your struct declaration is not packed, the layout of the input signal corresponding to the struct data type is determined by C language-specific padding rules that cause the Intel HLS Compiler to insert padding bytes before struct members that require a specific alignment.

  You should use packed structs as arguments to your RTL modules unless there is a specific reason to conform to a particular padded struct layout.
• **Packed Structs**
  If the struct type is declared as packed, member values start on an 8-bit boundary.

  The Intel HLS Compiler does not insert padding bytes to align struct members on platform-defined boundaries. The second-declared member always starts in the next highest byte after high-order byte of the first-declared struct member.

• **System Verilog Structs**
  If you are developing an RTL module in System Verilog, you can declare a System Verilog struct type that corresponds to the C++ struct type that is mapped to the input signal of your RTL module.

  The declaration order of the struct members is reversed in the System Verilog declaration because it specifies how the member signals should be concatenated to produce the composite signal. In a System Verilog concatenation expression, the bits are specified from high to low. That is, the last byte of the C++ struct type must be listed first in the System Verilog signal concatenation.

  You can compile your emulation models as HLS components to obtain an `interface_structs.v` file that contains declarations of the System Verilog struct types corresponding to the struct-type arguments of those functions. For details, see the following tutorial:

  `<quartus_installdir>/hls/examples/tutorials/libraries/rtl_struct_mapping`

• **Pointers in Structs**
  You cannot use struct types that have reference or pointer members as arguments to or return values from RTL-based functions.

**11.4.1.6. HLS Emulation Models for RTL-Based Functions**

For an RTL-based function, write C++ code that serves as an emulation model for that function. This model is used when you run your component in emulation mode.

The emulation model is not used when you simulate your component; simulations use RTL extracted from the library.

*Important:* If your function uses `static` variables to hold internal state, the emulation is equivalent to the RTL functionality only if the function is called from only one place in the HLS component.

This behavior is different because on CPUs all calls to the function share the same state variables. On FPGAs, the RTL module is instantiated once for each location in the HLS component where the function is called, and these instances do no share state.

**11.4.1.7. Potential Incompatibility between RTL Modules and Partial Reconfiguration**

When you create an HLS library using RTL modules, you might encounter Partial Reconfiguration-related issues.

Consider a situation where you create and verify your library on a device that does not support Partial Reconfiguration (PR). If you then use the library RTL module inside a PR region, the module might not function correctly after PR.

To ensure that the RTL modules function correctly on a device that uses PR:
• The RTL modules do not use memory logic array blocks (MLABs) with initialized content.
• The RTL modules do not make any assumptions regarding the power-up values of any logic.

For complete PR coding guidelines, refer to Creating a Partial Reconfiguration Design in the Partial Reconfiguration User Guide.

### 11.4.1.8. Stall-Free RTL

The Intel HLS Compiler can optimize hardware resource usage and performance by not placing stall logic around an RTL module with fixed latency.

If you have an RTL module with a fixed latency that you want integrated into your component pipeline without surrounding stall logic, ensure that you set attributes in the object manifest file (.xml) as follows:

1. Specify a value for the `EXPECTED_LATENCY` attribute (under the `FUNCTION` element) so that the latency equals the number of pipeline stages in the module.
   **Important:** An inaccurate `EXPECTED_LATENCY` value causes the RTL module to be out of sync with the rest of the pipeline, and can lead to functionally incorrect results.

2. Set the `IS_STALL_FREE` attribute under the `FUNCTION` element to "yes".

   This setting instructs the Intel HLS Compiler to avoid placing stall logic around the RTL module. This setting also tells the compiler that the RTL module produces a result after the number cycles specified in the `EXPECTED_LATENCY` attribute after accepting input values. The stall free logic produces a result every cycle but the result is delayed by the number cycles specified in the `EXPECTED_LATENCY` attribute.

For RTL modules with a fixed latency, the output signals (\texttt{ovvalid} and \texttt{oready}) can have constant high values, and the input ready signal (\texttt{iready}) can be ignored.

A stall-free RTL module might receive an invalid input signal (\texttt{invalid} is low). In this case, the module must produce invalid data on the output `EXPECTED_LATENCY` cycles after the cycle in which the input was invalid. For a stall-free RTL module without an internal state, you might find it convenient to propagate the invalid input through the module. If the module has an internal state, that state should not be affect by data inputs that are not accompanied by \texttt{invalid = 1}.

### 11.4.1.9. RTL Module Restrictions and Limitations for HLS Libraries

RTL modules that you want to include in an HLS library are subject to some restrictions and limitations to ensure that the library works consistently across different user designs.

**RTL Module Restrictions**

When you create an RTL module, ensure that it operates within the following restrictions:

• The RTL module must work correctly at any clock frequency that passes timing analysis.
• Data input and output sizes must match the sizes of the arguments and return value declared in the RTL module function signature (.h) file.
  
  For example, if you work with 24-bit values inside an RTL module, declare inputs to be 32 bits and declare the function signature to accept the `uint` data type. In the RTL module, accept the 32-bit input but discard the top 8 bits.

• RTL modules cannot connect to external I/O signals. All input and output signals must come from the HLS component that uses the library.

• An RTL module must have a `clock` port, a `resetn` port, and handshaking ports to support the data input and output interfaces. The handshake signal must be named `ivalid`, `ovalid`, `iready`, and `oready`.

• Every function call that corresponds to an RTL module instantiation is completely independent of other instantiations. No hardware is shared.

• An RTL module must receive all its inputs at the same time. A single `ivalid` input signifies that all inputs contain valid data.

**RTL-Based Object Limitations**

Using RTL modules in HLS libraries has the following limitations:

• You can only set RTL module parameters in the object manifest file (.xml) file.
  
  To use the same module with multiple parameter combinations, create a separate `FUNCTION` tag for each parameter combination.

• Pass data inputs to the RTL module only by value through the HLS component code.
  
  You cannot pass streams, pointers, or references as input to an RTL module. For streaming data, extract data from the stream first in your component and then pass the extracted scalar data to the RTL module in the HLS library.
  
  Passing data inputs to an RTL module as pointers or references causes a fatal error in the Intel HLS Compiler.

• Names of RTL module source files cannot conflict with the names of objects in other libraries or in file names of Intel HLS Compiler IP.
  
  When you create a library, choose RTL module names that are unlike to conflict with other libraries or compiler IP. For example, prefix the name of your RTL modules with the name of your library.
  
  If there is a naming conflict, the Intel Quartus Prime compilation of the HLS component might fail or result in a functionally-incorrect FPGA image.

• Names of the RTL module and its signals cannot conflict with reserved names defined by any of the supported RTL languages: Verilog, System Verilog, and VHDL.

• The Intel HLS Compiler does not support `.qip` files. You must manually parse nested `.qip` files to create a flat list of RTL files.
11.4.2. Creating a Static-Object File from an RTL Module

Before an RTL module can be included in a library intended for use in an Intel HLS Compiler design, create a platform-specific object (.o files on Linux, .obj files on Windows) from the RTL module. Use the `fpga_crossgen` command to create the object.

For instructions on creating an OpenCL library object file from RTL, see "Packaging an RTL Component for an OpenCL Library" in the Intel FPGA SDK for OpenCL Pro Edition Programming Guide. Before you can create an HLS library object from an RTL module, ensure that the functions in your RTL module are functionally correct and that you have the following files ready:

- RTL module source files
  These files are the Verilog (.v), System Verilog (.sv), or VHDL (.vhdl) files and the accompanying memory initialization files (.mif or .hex) that define the RTL modules.

- RTL object manifest file
  This XML file describes the callable interfaces of your RTL modules. Review Object Manifest File Syntax on page 124 for details about what to include in this XML file.

- HLS emulation model file
  These C++ files (.cpp and .h) provide an emulation model for the RTL module that allows you to emulate your component when it includes an HLS library that contains this RTL module. Full hardware compilations use the RTL source files.

- RTL module function signature file
  This C-style header file (.h) declares the signatures of the functions that are implemented by the RTL module and described in the object manifest file. Include this file in your HLS component code for the component to call the functions provided by the RTL modules packaged in the object.

1. After you have the files ready, create the HLS library object with the following command:

   ```
   fpga_crossgen <object_manifest_file_name> --target hls --emulation_model <emulation_model_file_location> [--o <object_file_name>]
   ```

   Where `<object_manifest_file_name>` is the full path of the RTL object manifest (.xml) file including the file name. This path can be a full or relative path.

   If you do not specify an object file name with the `-o` option, the object file name defaults to be the same name as the object manifest file name. That is, an object manifest file named `manifest.xml` results in an object file named `manifest.o` (on Linux) or `manifest.obj` (on Windows).

   The output of the command is a platform-specific object file ( (.o on Linux, .obj on Windows). The platform of the object file is determined by the platform where you run the `fpga_crossgen` command. When you run the command on Linux, you get a .o object file. When you run the command on Windows, you get a .obj object file.

   **Important:** Each object created with the `fpga_crossgen` command is assigned a compiler version number. You can package only object with the same version number into a library, and a library can be used only with a target compiler (For example, i++) with the same version number.
11.5. Packaging Object Files Into a Library

Collect object files in a library file so that others can incorporate the library into their projects and call the functions that are contained in the objects in the library. Package object files into a library with the fpga_libtool command.

Before you package object files into a library, ensure that you have the path information for all of the object files that you want to include in the library.

All objects to be packaged in the library must have the same version number. This library can be used only by an Intel high-level design tool with the same version number.

The fpga_libtool command creates libraries encapsulated in operating system specific archive files (.a on Linux, .lib on Windows). Create the HLS library file with the following command:

```
fpga_libtool --target target_HLD_tool --create library_name [.a | .lib | .aoclib] object_file_1 [object_file_2 ... object_file_n]
```

Where the command parameters are defined as follows:

- **target_HLD_tool**
  The target Intel high-level design tool for this library. This parameter can have one of the following values:
  
  - **hls**
    Target this library for components developed with the Intel HLS Compiler.
    Libraries built for the Intel HLS Compiler are encapsulated in operating system specific archive files (.a on Linux, .lib on Windows). You cannot use HLS libraries created on one operating system with the Intel HLS Compiler running on a different operating system.
  
  - **aoc**
    Target this library for kernels developed with the Intel FPGA SDK for OpenCL.
    Libraries built for the Intel FPGA SDK for OpenCL are not operating system specific. The objects are created as Intel FPGA SDK for OpenCL object files (.aoclib).
    You must have the Intel FPGA SDK for OpenCL Pro Edition installed to use this option. The version of the SDK must be the same as your version of Intel HLS Compiler.
- **library_name**
  
  The name of the library file.

  Specify the file extension of the library files as follows, depending on the target high-level design tool:

  - Intel HLS Compiler
    
    Specify the operating-system specific archive extension: .a for Linux-platform libraries and .lib for Windows-platform libraries.

  - Intel FPGA SDK for OpenCL
    
    Specify .aoclib as the file extension for an OpenCL library.

    OpenCL libraries are not operating-system specific.

You can specify one or more object files to include in the library.

For example, the following command packages three Linux-platform objects (prim1.o, prim2.o, and prim3.o) into an HLS library called libdemo:

```
fpga_libtool --create libdemo.a prim1.o prim2.o prim3.o --target hls
```
12. Advanced Hardware Synthesis Controls

12.1. Hardware Implementation Control for Math Functions

For math functions performed on certain data types, you can control whether the function is implemented in hardware using DSP blocks or soft logic (ALMs).

You can review the math function hardware implementation by reviewing the details of the instruction nodes for the math function in System Viewer of the High-Level Design Reports (report.html).

For an example of controlling the hardware implementation of math functions, refer to the following tutorial:

<quartus_installdir>/hls/examples/tutorials/best_practices/control_of_dsp_usage

**Supported Data Types and Math Functions**

Refer to Math Function Hardware Implementation Summary on page 140.

**Global-Scope Control**

You can control the hardware implementation of supported math functions at a global scope for your design by specifying the --dsp-mode option of the i++ command.

If you use both global-scope control and local-scope control, any local-scope controls defined in your component or tasks override the global-scope control setting.

For details, refer to The --dsp-mode Command Option on page 141.
Local-Scope Control

You can control the hardware implementation of supported math functions at a local scope with the `ihc::math_dsp_control` function. This function is defined in the `HLS/math_dsp_control.h` header file.

If you use both local-scope control and global-scope control, any local-scope controls defined in your component or tasks override the global-scope control setting.

For details, refer to The `ihc::math_dsp_control` Function on page 141.

### 12.1.1. Math Function Hardware Implementation Summary

The following table summarizes the default hardware implementations for math functions for different data type and what combinations of data type and math function support hardware implementation control.

This table applies to the following device families: Intel Agilex, Intel Arria 10, and Intel Stratix 10.

**Table 34. Math Function Hardware Implementation Summary**

This table does not list operations between two constants because they are always implemented by soft-logic, and are often optimized away by the compiler.

<table>
<thead>
<tr>
<th>Operand 1 Data Type</th>
<th>Operand 2 Data Type</th>
<th>Addition</th>
<th>Subtraction</th>
<th>Multiplication</th>
<th>Division</th>
<th>Dot Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>float Variable</td>
<td>float Variable</td>
<td>(D)/S</td>
<td>(D)/S</td>
<td>D</td>
<td>S+D</td>
<td>D</td>
</tr>
<tr>
<td>float Constant</td>
<td>(D)/S</td>
<td></td>
<td></td>
<td>S</td>
<td></td>
<td>D</td>
</tr>
<tr>
<td>double Variable</td>
<td>double Variable</td>
<td>S</td>
<td>S</td>
<td>S+D</td>
<td>S+D</td>
<td>S+D</td>
</tr>
<tr>
<td>double Constant</td>
<td>S</td>
<td></td>
<td></td>
<td>S</td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>int/ac_int Variable</td>
<td>int/ac_int Variable</td>
<td>S</td>
<td>S</td>
<td>S+D</td>
<td>S</td>
<td>S+D</td>
</tr>
<tr>
<td>int/ac_int Constant</td>
<td>S</td>
<td></td>
<td></td>
<td>S</td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>ac_fixed Variable</td>
<td>ac_fixed Variable</td>
<td>S</td>
<td>S</td>
<td>S+D</td>
<td>S+D</td>
<td>S+D</td>
</tr>
<tr>
<td>ac_fixed Constant</td>
<td>S</td>
<td></td>
<td></td>
<td>S</td>
<td></td>
<td>S</td>
</tr>
<tr>
<td>hls_float Variable</td>
<td>hls_float Variable</td>
<td>The <code>hls_float&lt;8,23&gt;</code> data type uses the same implementation as float. Other <code>hls_float</code> data types use the same implementation as double.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hls_float Constant</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

S  Math function is implemented only in soft logic with ALMs.
**12. Advanced Hardware Synthesis Controls**

Math function is implemented only using DSPs.

\[ D \]

Math function is implemented with a combination of soft logic (ALMs) and DSPs.

\[ S+D \]

Bold text indicates that hardware implementation controls are available for that data type and math function.

\[ (D)/S \]

Bold-italic text in parentheses indicates the default hardware implementation in the absence of applying any hardware implementation controls.

### 12.1.2. The `--dsp-mode` Command Option

Use the `--dsp_mode` option of the `i++` command to control the hardware implementation for supported data types and math functions for your entire compilation. For local-scope control, use `ihc::math_dsp_control` function.

Specify the `--dsp-mode` command option as follows:

```
i++ --dsp-mode={prefer-dsp | prefer-softlogic | default}
```

**Table 35. `--dsp-mode` Command Option Values**

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>prefer-dsp</td>
<td>The compiler tries to implement supported math operations with DSP blocks.</td>
</tr>
<tr>
<td>prefer-softlogic</td>
<td>The compiler tries to implement supported math operations with soft logic using ALMs.</td>
</tr>
<tr>
<td>default</td>
<td>The compiler implements supported math operations with DSP blocks or soft logic (ALMs) based on data type and operation. This value is the default.</td>
</tr>
</tbody>
</table>

### 12.1.3. The `ihc::math_dsp_control` Function

Use the `ihc::math_dsp_control` function to control the hardware implementation for supported data types and math functions at a local-scope.

The `ihc::math_dsp_control` function is defined in the HLS/math_dsp_control.h header file. Include this header file in your component or task code to use the function.

**Syntax**

```
ihc::math_dsp_control<HW_impl_pref, lambda_expr_propagation>
```

where the template parameters are defined as follows:

**HW_impl_pref**

The `HW_impl_pref` template parameter takes an `enum` data type with one of the following values:
Table 36. HW_impl_pref Values

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::Preference::DSP</td>
<td>The compiler tries to implement supported math operations with DSP blocks.</td>
</tr>
<tr>
<td></td>
<td>This value is the default.</td>
</tr>
<tr>
<td>ihc::Preference::Softlogic</td>
<td>The compiler tries to implement supported math operations with soft logic</td>
</tr>
<tr>
<td></td>
<td>using ALMs.</td>
</tr>
<tr>
<td>ihc::Preference::Compiler_default</td>
<td>The compiler implements supported math operations with DSP blocks or soft</td>
</tr>
<tr>
<td></td>
<td>logic (ALMs) based on data type and operation.</td>
</tr>
</tbody>
</table>

**lambda_expr_propagation**

The `lambda_expr_propagation` template parameter is a Boolean value that determines the propagation of the `HW_impl_pref` value to function calls in the lambda expression as follows:

Table 37. lambda_expr_propagation Values

<table>
<thead>
<tr>
<th>Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>true</td>
<td>The <code>HW_impl_pref</code> value is propagated to function calls in the lambda expression function. This value is the default.</td>
</tr>
<tr>
<td>false</td>
<td>The <code>HW_impl_pref</code> value is not propagated to function calls in the lambda expression function. The <code>HW_impl_pref</code> value applies only to math operations directly called in the lambda expression.</td>
</tr>
</tbody>
</table>

**Nested Calls**

A `HW_impl_pref` value never applies to any nested `ihc::math_dsp_control` function calls. Each `ihc::math_dsp_control` function call requires its own `HW_impl_pref` value specified.

**Example**

```c
component float test_lib(float a, float b) {
    ihc::math_dsp_control<ihc::Preference::Softlogic>({[
        a += 7.89f; // Soft-logic
        a += b;     // Soft-logic
    ]});
    return a;
}
```

12.2. The `hls_fpga_reg()` Function

In some cases, explicitly asking the compiler to insert a register stage between the operand and the return value of the function call can help improve the performance of your component. Use the `hls_fpga_reg()` function to insert at least one register between the operand and return value of the function call.

Typically, you do not need to use this function to achieve the performance from your component that you want.

To use the `hls_fpga_reg()` function effectively, you must know about how portions of the data path are placed on FPGA devices, and you typically use the `hls_fpga_reg()` function for the following purposes:
• Breaking the critical paths between spatially distant portions of a data path, such as between processing elements of a large systolic array.

• Reducing the pressure on placement and routing efforts caused by spatially distinct portions of the component implementation.

The Intel HLS Compiler does not provide feedback about where you should add `hls_fpga_reg()` function calls. Use Intel Quartus Prime software to determine where you should insert the calls to address specific aspects of component performance.

You can learn more about the `hls_fpga_reg()` function by reviewing the tutorial available in the following location:

<quartus_installdir>/hls/examples/tutorials/best_practices/fpga_reg

**Syntax**

```
T hls_fpga_reg(T op)
```

where `T` can be any sized type.

**Description**

The `hls_fpga_reg()` function directs the Intel HLS Compiler to insert at least one hardware pipelining register on the signal path that assigns the operand to the return value. This built-in function operates as an assignment, where the operand is assigned to the return value. The assignment has no implicit semantic or functional meaning beyond a standard assignment.

**Usage Notes**

You can nest `hls_fpga_reg()` function calls to increase the minimum number of registers that are inserted on the assignment path. Because each function call guarantees the insertion of at least one register stage, the number of nested calls provides a lower limit on the number of registers.

For example, the following code snippet tells the compiler to insert at least two registers on the assignment path.

```c
int out=hls_fpga_reg(hls_fpga_reg(in));
```

The compiler inserts two or more registers on the path.
13. Intel High Level Synthesis Compiler Pro Edition

Reference Summary

13.1. Intel HLS Compiler Pro Edition i++ Command-Line Arguments

Use the i++ command-line arguments to affect how your component is compiled and linked.

General i++ Command Options

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>--debug-log</td>
<td>Generate the compiler diagnostics log.</td>
</tr>
<tr>
<td>-h, --help</td>
<td>List compiler command options along with brief descriptions.</td>
</tr>
<tr>
<td>-o result</td>
<td>Place compiler output into the &lt;result&gt; executable and the &lt;result&gt;.prj directory.</td>
</tr>
<tr>
<td>-v</td>
<td>Display messages describing the progress of the compilation.</td>
</tr>
<tr>
<td>--version</td>
<td>Display compiler version information.</td>
</tr>
</tbody>
</table>

Command Options Affecting Compiling

<table>
<thead>
<tr>
<th>Option</th>
<th>Default Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-c</td>
<td></td>
<td>Preprocess, parse, and generate object files.</td>
</tr>
<tr>
<td>--component</td>
<td></td>
<td>Comma-separated list of function names to compile to RTL.</td>
</tr>
<tr>
<td>component name</td>
<td></td>
<td>To use this option, your component must be configured with C-linkage using the extern &quot;C&quot; specification. For example: extern &quot;C&quot; int myComponent(int a, int b) Using the component function attribute is preferred over using the --component command option to indicate functions that you want the compile to RTL.</td>
</tr>
<tr>
<td>-Dmacro [=val]</td>
<td></td>
<td>Define a &lt;macro&gt; with &lt;val&gt; as its value.</td>
</tr>
<tr>
<td>-g</td>
<td></td>
<td>Generate debug information (default option).</td>
</tr>
<tr>
<td>-g0</td>
<td></td>
<td>Do not generate debug information.</td>
</tr>
<tr>
<td>--gcc-toolchain=&lt;GCC_dir&gt;</td>
<td></td>
<td>Specifies the path to a GCC installation that you want to use for compilation. This path should be the absolute path to the directory that contains the GCC lib, bin, and include folders.</td>
</tr>
<tr>
<td>--hyper-optimized-handshaking=[auto</td>
<td>off]</td>
<td>auto</td>
</tr>
</tbody>
</table>

continued...
<table>
<thead>
<tr>
<th>Option</th>
<th>Default Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-I</td>
<td>&lt;dir&gt;</td>
<td>Add directory <code>&lt;dir&gt;</code> to the end of the main include path.</td>
</tr>
<tr>
<td>-march=</td>
<td>x86-64</td>
<td>Generate code for an emulator flow (x86-64) or for the specified FPGA family or FPGA part number.</td>
</tr>
<tr>
<td>--quartus-compile</td>
<td></td>
<td>Run the HDL generated through Intel Quartus Prime to generate accurate fMAX and area estimates. Your component is not expected to cleanly close timing.</td>
</tr>
<tr>
<td>--quartus-seed &lt;seed&gt;</td>
<td></td>
<td>Specifies the Fitter seed to use when your component is compiled to hardware by Intel Quartus Prime.</td>
</tr>
</tbody>
</table>
| --simulator <simulator_name> | modelsim | Specifies the simulator you are using to perform verification. This command option can take the following values for `<simulator_name>`:  
  - `modelsim` Use ModelSim for component verification.  
  - `none` Disable verification. That is, generate RTL for components without the test bench.  
  If you do not specify this option, `--simulator modelsim` is assumed. |
| -ffp-contract=fast |           | Remove intermediate rounding and conversion when possible, except for code blocks fenced by `#pragma clang fp contract(off)`.  
To learn more, review the following tutorial: `<quartus_installdir>/hls/examples/tutorials/best_practices/floating_point_ops` |
| -ffp-reassoc | Deprecated: This command option is deprecated and will be removed in a later release of the Intel HLS Compiler. Use the `-ffp-reassociate` command option instead.  
Relax the order of floating point arithmetic operations, except for code blocks fenced by `#pragma clang fp reassoc(off)`  
To learn more, review the following tutorial: `<quartus_installdir>/hls/examples/tutorials/best_practices/floating_point_ops` |
| -ffp-reassociate |             | Relax the order of floating point arithmetic operations, except for code blocks fenced by `#pragma clang fp reassociate(off)`  
To learn more, review the following tutorial: `<quartus_installdir>/hls/examples/tutorials/best_practices/floating_point_ops` |
| --daz | For double data types only, disable subnormal support in double-precision floating-point computations. |
### Option --rounding=[ieee | faithful]

<table>
<thead>
<tr>
<th>Default Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>For double data types only, control rounding scheme for double-precision adders, multipliers, and dividers. If you do not specify this option, adders and multipliers use IEEE-754 round to nearest, ties to even (RNE) rounding (0.5 ULP) and dividers uses faithful rounding (1 ULP). The --rounding option can take one of the following values:</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>ieee</td>
<td>All adders, multipliers, and dividers use IEEE-754 RNE rounding. IEEE-754 RNE rounding rounds results to the nearest value. If the number falls midway, it is rounded to the nearest value with an even least-significant digit. This is the default rounding mode defined by IEEE 754-2008 standard.</td>
</tr>
<tr>
<td>faithful</td>
<td>All adders, multipliers, and dividers use faithful rounding. Faithful rounding rounds results to either the upper or lower nearest single-precision numbers. Therefore, faithful rounding produces one of two possible values. The choice between the two is not defined. Faithful rounding has a maximum error of one ULP. Errors are not guaranteed to be evenly distributed. Faithful rounding mode is not defined by the IEEE-754 standard.</td>
</tr>
</tbody>
</table>

### Option --clock clock target

<table>
<thead>
<tr>
<th>Default Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>240 MHz</td>
<td>Optimize the RTL for the specified clock frequency or period. The clock target value must include a unit. For example:</td>
</tr>
<tr>
<td></td>
<td>i++ -march=&quot;Arria 10&quot; test.cpp --clock 100MHz</td>
</tr>
<tr>
<td></td>
<td>i++ -march=&quot;Arria 10&quot; test.cpp --clock 10ns</td>
</tr>
</tbody>
</table>

### Option --dsp_mode=[prefer-dsp | prefer-softlogic | default]

<table>
<thead>
<tr>
<th>Default Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>default</td>
<td>For supported data types and math operations, controls the hardware implementation of math functions on a global scope.</td>
</tr>
<tr>
<td>prefer-dsp</td>
<td>The compiler tries to implement supported math operations with DSPs.</td>
</tr>
<tr>
<td>prefer-softlogic</td>
<td>The compiler tries to implement supported math operations with soft logic using ALMs.</td>
</tr>
<tr>
<td>default</td>
<td>The compiler implements supported math operations with DSP blocks or soft logic (ALMs) based on data type and operation. This value is the default.</td>
</tr>
</tbody>
</table>

For details about the implementation of math functions in hardware, refer to Math Function Hardware Implementation Summary on page 140. To learn more, refer to the following tutorial: <quartus_installdir>/hls/examples/tutorials/best_practices/control_of_dsp_usage

---

### Command Options Affecting Linking

<table>
<thead>
<tr>
<th>Option</th>
<th>Default Value</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>-ghdl[=&lt;depth&gt;]</td>
<td>Enable full debug visibility and logging of HDL signals in simulation. Use the optional &lt;depth&gt; attribute to specify how many levels of hierarchy are logged. If you do not specify a value for the &lt;depth&gt; attribute, all signals are logged.</td>
<td></td>
</tr>
</tbody>
</table>

---

*continued...*
### 13.2. Intel HLS Compiler Pro Edition Header Files

Coding your component to be compiled by the Intel HLS Compiler requires you to include the `hls.h` header file. Other header files provided with the Intel HLS Compiler provide FPGA-optimized implementations of certain C and C++ functions.

<table>
<thead>
<tr>
<th>HLS Header File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HLS/hls.h</td>
<td>Required for component identification and component parameter interfaces.</td>
</tr>
<tr>
<td>HLS/math.h</td>
<td>Includes FPGA-specific definitions for the math functions from the <code>math.h</code> for your operating system.</td>
</tr>
<tr>
<td>HLS/extendedmath.h</td>
<td>Includes additional FPGA-specific definitions of math functions not in <code>math.h</code>.</td>
</tr>
<tr>
<td>HLS/ac_int.h</td>
<td>Provides FPGA-optimized arbitrary width integer support.</td>
</tr>
<tr>
<td>HLS/ac_fixed.h</td>
<td>Provides FPGA-optimized arbitrary precision fixed point support.</td>
</tr>
<tr>
<td>HLS/ac_fixed_math.h</td>
<td>Provides FPGA-optimized arbitrary precision fixed point math functions.</td>
</tr>
<tr>
<td>HLS/ac_complex.h</td>
<td>Provides FPGA-optimized complex number support.</td>
</tr>
<tr>
<td>HLS/hls_float.h</td>
<td>Provides FPGA-optimized arbitrary-precision IEEE-754 compliant floating-point number support.</td>
</tr>
<tr>
<td>HLS/hls_float_math.h</td>
<td>Provides FPGA-optimized floating-point math functions.</td>
</tr>
<tr>
<td>HLS/math_dsp_control.h</td>
<td>Provides the <code>ihc::math_dsp_control</code> function to control the hardware implementation for supported data types and math functions at a local-scope.</td>
</tr>
<tr>
<td>HLS/stdio.h</td>
<td>Provides <code>printf</code> support for components so that <code>printf</code> statements work in x86 emulations, but are disabled in component when compiling to an FPGA architecture.</td>
</tr>
<tr>
<td><code>&lt;iostream&gt;</code></td>
<td>To use <code>cout</code> and <code>cerr</code> in your component, guard the statements with the <code>HLS_SYNTHESIS</code> macro.</td>
</tr>
</tbody>
</table>

### hls.h Header File

**Syntax**  
```
#include "HLS/hls.h"
```

**Description**  
Required for component identification and component parameter interfaces.
**math.h** Header File

**Syntax**
```
#include "HLS/math.h"
```

**Description**
Includes FPGA-specific definitions for the math functions from the `math.h` for your operating system.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/best_practices/single_vs_double_precision_math.

**extendedmath.h** Header File

**Syntax**
```
#include "HLS/extendedmath.h"
```

**Description**
Includes additional FPGA-specific definitions of math functions not in `math.h`.

To learn more, review the following design:
<quartus_installdir>/hls/examples/QRD.

**ac_int.h** Header File

**Syntax**
```
#include "HLS/ac_int.h"
```

**Description**
Intel HLS Compiler version of `ac_int` header file.
Provides FPGA-optimized arbitrary width integer support.

To learn more, review the following tutorials:
- <quartus_installdir>/hls/examples/tutorials/ac_datatypes/ac_int_basic_ops
- <quartus_installdir>/hls/examples/tutorials/ac_datatypes/ac_int_overflow
- <quartus_installdir>/hls/examples/tutorials/best_practices/struct_interfaces

**ac_fixed.h** Header File

**Syntax**
```
#include "HLS/ac_fixed.h"
```

**Description**
Intel HLS Compiler version of the `ac_fixed` header file.
Provides FPGA-optimized arbitrary precision fixed point support.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/ac_datatypes/ac_fixed_constructor.
ac_fixed_math.h Header File

Syntax #include "HLS/ac_fixed_math.h"

Description Intel HLS Compiler version of the ac_fixed_math header file. Provides FPGA-optimized arbitrary precision fixed point math functions.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/ac_datatypes/ac_fixed_math_library.

ac_complex.h Header File

Syntax #include "HLS/ac_complex.h"

Description Intel HLS Compiler version of the ac_complex header file. Provides FPGA-optimized complex math functions.

hls_float.h Header File

Syntax #include "HLS/hls_float.h"

Description Header file to provide FPGA-optimized arbitrary-precision IEEE 754 compliant floating-point number support.

To learn more, review the following tutorials:
• <quartus_installdir>/hls/examples/tutorials/hls_float/1_reduced_double
• <quartus_installdir>/hls/examples/tutorials/hls_float/2_explicit_arithmetic
• <quartus_installdir>/hls/examples/tutorials/hls_float/3_conversions

hls_float_math.h Header File

Syntax #include "HLS/hls_float_math.h"

Description Header file to provide math functions for hls_float data types.

To learn more, review the following tutorials:
math_dsp_control.h Header File

Syntax #include "HLS/math_dsp_control.h"

Description Header file to provide the ihc::math_dsp_control function to control the hardware implementation for supported data types and math functions at a local-scope.

For control at a global-scope, use the --dsp-mode option of the i++ command.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/best_practices/control_of_dsp_usage

stdio.h Header File

Syntax #include "HLS/stdio.h"

Description Provides printf support for components so that printf statements work in x86 emulations, but are disabled in component when compiling to an FPGA architecture.

Standard C++ <iostream> Header File

Syntax #include <iostream>

Description To use the C++ standard output streams (cout and cerr) provided by the standard <iostream> header, you must guard any standard output statements with the HLS_SYNTHESIS macro.

This macro ensures that statements in a component work in x86 emulations but are disabled in the component when compiling to an FPGA architecture.

13.3. Intel HLS Compiler Pro Edition Compiler-Defined Preprocessor Macros

The Intel HLS Compiler Pro Edition has built-in macros that you can use to customize your code to create flow-dependent behaviors.
Table 39. Macro Definition for __INTELFPGA_COMPILER__

<table>
<thead>
<tr>
<th>Tool Invocation</th>
<th><strong>INTELFPGA_COMPILER</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>g++ or cl</td>
<td>Undefined</td>
</tr>
<tr>
<td>i++ -march=x86-64</td>
<td>2120</td>
</tr>
<tr>
<td>i++ -march=&quot;&lt;FPGA_family_or_part_number&gt;&quot;</td>
<td>2120</td>
</tr>
</tbody>
</table>

Table 40. Macro Definition for HLS_SYNTHESIS

<table>
<thead>
<tr>
<th>Tool Invocation</th>
<th>HLS_SYNTHESIS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Testbench Code</td>
</tr>
<tr>
<td>g++ or cl</td>
<td>Undefined</td>
</tr>
<tr>
<td>i++ -march=x86-64</td>
<td>Undefined</td>
</tr>
<tr>
<td>i++ -march=&quot;&lt;FPGA_family_or_part_number&gt;&quot;</td>
<td>Undefined</td>
</tr>
</tbody>
</table>

13.4. Intel HLS Compiler Pro Edition Keywords

Table 41. Intel HLS Compiler Pro Edition Keywords

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>component</td>
<td>Indicates that a function is a component. Example:</td>
</tr>
<tr>
<td></td>
<td>component void foo()</td>
</tr>
</tbody>
</table>

13.5. Intel HLS Compiler Pro Edition Simulation API (Testbench Only)

Table 42. Intel HLS Compiler Pro Edition Simulation API (Testbench only) Summary

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc_hls_enqueue</td>
<td>This function enqueues one invocation of an HLS component.</td>
</tr>
<tr>
<td>ihc_hls_enqueue_noret</td>
<td>This function enqueues one invocation of an HLS component. This function should be used when the return type of the HLS component is void.</td>
</tr>
<tr>
<td>ihc_hls_component_run_all</td>
<td>This function pushes all enqueued invocations of a component into the component in the HDL simulator as quickly as the component can accept new invocations.</td>
</tr>
<tr>
<td>ihc_hls_sim_reset</td>
<td>This function sends a reset signal to the component during automated simulation.</td>
</tr>
<tr>
<td>ihc_hls_set_component_wait_cycle</td>
<td>This function tells the simulation process to continue running for a specified number of additional cycles (beyond the default wait period of 100 cycles) after the done signal for the specified component is observed.</td>
</tr>
</tbody>
</table>
ihc_hls_enqueue Function

Syntax

ihc_hls_enqueue(void* retptr, void* funcptr, /*function arguments*/)

Description

This function enqueues one invocation of an HLS component. The return value is stored in the first argument which should be a pointer to the return type. The component is not run until the ihc_hls_component_run_all() is invoked.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/usability/enqueue_call`.

ihc_hls_enqueue_noret Function

Syntax

ihc_hls_enqueue_noret(void* funcptr, /*function arguments*/)

Description

This function enqueues one invocation of an HLS component. This function should be used when the return type of the HLS component is void. The component is not run until the ihc_hls_component_run_all() is invoked.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/usability/enqueue_call`.

ihc_hls_component_run_all Function

Syntax

ihc_hls_component_run_all (void* funcptr)

Description

This function accepts a pointer to the HLS component function. When run, all enqueued invocations of the component will be pushed into the component in the HDL simulator as quickly as the component can accept new invocations.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/usability/enqueue_call`.

ihc_hls_sim_reset Function

Syntax

int ihc_hls_sim_reset (void)

Description

This function sends a reset signal to the component during automated simulation. It returns 1 if the reset was exercised or 0 otherwise.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/component_memories/static_var_init`. 
**ihc_hls_set_component_wait_cycle Function**

**Syntax**

```
ihc_hls_set_component_wait_cycle(<component function name>, <# of wait cycles>)
```

**Description**

This function tells the simulation process to continue running for a specified number of additional cycles (beyond the default wait period of 100 cycles) after the `done` signal for the specified component is observed. This delay can enable task functions with a higher latency than the component function to successfully return their output during simulation.

Use this function when you simulate a design that uses a system of tasks where the completion of a task function is not synchronized with an `ihc::collect` call.

By default, the simulation process simulates an additional 100 cycles after a component asserts the `done` signal to ensure all operations have propagated back to the testbench. This function tells the simulation process for the specified component to continue running for the specified number of cycles in addition to the default wait period of 100 cycles.

**Simulation API Code Example**

```c
component int foo(int val) {
    // function definition
}

component void bar (int val) {
    // function definition
}

int main() {
    // …….
    int input = 0;
    int res[5];
    ihc_hls_enqueue(&res, &foo, input);
    ihc_hls_enqueue_noret(&bar, input);
    input = 1;
    ihc_hls_enqueue(&res, &foo, input);
    ihc_hls_enqueue_noret(&bar, input);
    ihc_hls_component_run_all(&foo);
    ihc_hls_component_run_all(&bar);
}
```

**13.6. Intel HLS Compiler Pro Edition Component Memory Attributes**

Use the component memory attributes to control the on-chip component memory architecture of your component.

**Table 43. Intel HLS Compiler Pro Edition Component Memory Attributes Summary**

<table>
<thead>
<tr>
<th>Memory Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>hls_force_pow2_depth</code></td>
<td>Specifies that the memory implementing the variable or array has power-of-2 depth.</td>
</tr>
<tr>
<td><code>hls_register</code></td>
<td>Forces a variable or array to be carried through the pipeline in registers.</td>
</tr>
</tbody>
</table>

*continued...*
<table>
<thead>
<tr>
<th>Memory Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>hls_memory</strong></td>
<td>Forces a variable or array to be implemented as embedded memory.</td>
</tr>
<tr>
<td><strong>hls_memory_impl</strong></td>
<td>Forces a variable or array to be implemented as embedded memory of a specified type.</td>
</tr>
<tr>
<td><strong>hls_singlepump</strong></td>
<td>Specifies that the memory implementing the variable or array must be clocked at the same rate as the component accessing the memory.</td>
</tr>
<tr>
<td><strong>hls_doublepump</strong></td>
<td>Specifies that the memory implementing the variable or array must be clocked at twice the rate as the component accessing the memory.</td>
</tr>
<tr>
<td><strong>hls_numbanks</strong></td>
<td>Specifies that the memory implementing the variable or array must have a defined number of memory banks.</td>
</tr>
<tr>
<td><strong>hls_bankwidth</strong></td>
<td>Specifies that the memory implementing the variable or array must have memory banks of a defined width.</td>
</tr>
<tr>
<td><strong>hls_bankbits</strong></td>
<td>Forces the memory system to split into a defined number of memory banks and defines the bits used to select a memory bank.</td>
</tr>
<tr>
<td><strong>hls_simple_dual_port_memory</strong></td>
<td>Specifies that the memory implementing the variable or array should have no port that services both reads and writes.</td>
</tr>
<tr>
<td><strong>hls_merge (depthwise)</strong></td>
<td>Allows merging two or more local variables to be implemented in component memory as a single merged memory system in a depth-wise manner.</td>
</tr>
<tr>
<td><strong>hls_merge (widthwise)</strong></td>
<td>Allows merging two or more local variables to be implemented in component memory as a single merged memory system in a width-wise manner.</td>
</tr>
<tr>
<td><strong>hls_init_on_reset</strong></td>
<td>Forces the static variables inside the component to be initialized when the component reset signal is asserted.</td>
</tr>
<tr>
<td><strong>hls_init_on_powerup</strong></td>
<td>Sets the component memory implementing the static variable to initialize on power-up when the FPGA is programmed.</td>
</tr>
<tr>
<td><strong>hls_max_concurrency</strong></td>
<td>Deprecated: This attribute is deprecated and will be removed in a future release. Use the <strong>hls_private_copies</strong> memory attribute instead. Specifies that the memory implementing the variable or array has a defined number of private copies to allow concurrent iterations of a loop at any given time.</td>
</tr>
<tr>
<td><strong>hls_max_replicates</strong></td>
<td>Specifies that the memory implementing the variable or array has no more than the specified number of replicates to enable simultaneous reads from the datapath</td>
</tr>
<tr>
<td><strong>hls_private_copies</strong></td>
<td>Specifies that the memory implementing the variable or array has a defined number of private copies to allow concurrent iterations of a loop at any given time.</td>
</tr>
</tbody>
</table>

**hls_force_pow2_depth Memory Attribute**

**Syntax**

```
hls_force_pow2_depth(N)
```

**Constraints**

- \( N \) can be only 0 or 1.

**Default Value**

1
Specifies that the memory implementing the variable or array has a power-of-2 depth. This option is enabled if \( N = 1 \), and disabled if \( N = 0 \).

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/component_memories/non_power_of_two_memory.

### hls_register Memory Attribute

**Syntax**

hls_register

**Constraints**

N/A

**Default Value**

Based on the memory access pattern inferred by the compiler.

**Description**

Forces a variable or array to be implemented as registers.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/best_practices/swap_vs_copy.

### hls_memory Memory Attribute

**Syntax**

hls_memory

**Constraints**

N/A

**Default Value**

Based on the memory access pattern inferred by the compiler.

**Description**

Forces a variable or array to be implemented as embedded memory.

To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_implementation.

### hls_memory_impl Memory Attribute

**Syntax**

hls_memory_impl("type")

**Constraints**

N/A

**Default Value**

Based on the memory size and memory access pattern inferred by the compiler.

**Description**

Forces a variable or array to be implemented as embedded memory of the specified type.

The type parameter can be one of the following values:
**BLOCK_RAM**  Implement the variable or array as memory blocks, such as M20K memory blocks.

**MLAB**  Implement the variable or array as memory logic array blocks (MLABs).

To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_implementation.

---

**hls_singlepump** Memory Attribute

**Syntax**
hls_singlepump

**Constraints**  N/A

**Default Value**  Based on the memory access pattern inferred by the compiler.

**Description**  Specifies that the memory implementing the variable or array must be clocked at the same rate as the component accessing the memory.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/QRD.

---

**hls_doublepump** Memory Attribute

**Syntax**
hls_doublepump

**Constraints**  N/A

**Default Value**  Based on the memory access pattern inferred by the compiler.

**Description**  Specifies that the memory implementing the variable or array must be clocked at twice the rate of the component accessing the memory.

To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_bank_configuration.

---

**hls_numbanks** Memory Attribute

**Syntax**
hls_numbanks(N)

**Constraints**  This attribute is subject to constraints outlined in Constraints on Attributes for Memory Banks on page 57.

**Default Value**  Based on the memory access pattern inferred by the compiler.
**Description**

Specifies that the memory implementing the variable or array must have \( N \) banks, where \( N \) is a power-of-two constant number.

To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_geometry.

**hls_bankwidth Memory Attribute**

**Syntax**

\[ hls\_bankwidth(\ N) \]

**Constraints**

This attribute is subject to constraints outlined in Constraints on Attributes for Memory Banks on page 57.

**Default Value**

Based on the memory access pattern inferred by the compiler.

**Description**

Specifies that the memory implementing the variable or array must have banks that are \( N \) bytes wide, where \( N \) is a power-of-two constant number.

To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_geometry.

**hls_bankbits Memory Attribute**

**Syntax**

\[ hls\_bankbits(b_0, \ b_1, \ldots, \ b_n) \]

**Constraints**

This attribute is subject to constraints outlined in Constraints on Attributes for Memory Banks on page 57.

**Default Value**

Based on the memory access pattern inferred by the compiler.

**Description**

Forces the memory system to split into \( 2^{n+1} \) banks, with \( \{b_0, b_1, \ldots, b_n\} \) forming the bank-select bits.

**Important:** \( b_0, b_1, \ldots, b_n \) must be consecutive, positive integers. You can specify the consecutive, positive integers in ascending or descending order.

If you do not specify the \( hls\_bankwidth(\ N) \) attribute along with this attribute, then \( b_0, b_1, \ldots, b_n \) are mapped to array index bits 0 to \( n-1 \) in the memory bank implementation.

To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_geometry.
**hls_simple_dual_port_memory Memory Attribute**

**Syntax**

```
hls_simple_dual_port_memory
```

**Constraints**

N/A

**Default Value**

N/A

**Description**

Specifies that the memory implementing the variable or array should have no port that services both reads and writes.

To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_bank_configuration.

**hls_merge (depthwise) Memory Attribute**

**Syntax**

```
hls_merge("mem_name", "depth")
```

**Constraints**

N/A

**Default Value**

N/A

**Description**

Allows merging two or more local variables to be implemented in component memory as a single merged memory system in a depth-wise manner.

All variables with same `mem_name` label specified in their hls_merge attribute are merged into the same memory system.

To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_merging.

**hls_merge (widthwise) Memory Attribute**

**Syntax**

```
hls_merge("mem_name", "width")
```

**Constraints**

N/A

**Default Value**

N/A

**Description**

Allows merging two or more local variables to be implemented in component memory as a single merged memory system in a width-wise manner.

All variables with same `mem_name` label specified in their hls_merge attribute are merged into the same memory system.
To learn more, review the following tutorial:
<quartus_installdir>/hls/tutorials/component_memories/memory_merging.

**hls_init_on_reset** Memory Attribute

**Syntax**

hls_init_on_reset

**Constraints**

N/A

**Default Value**

Default behavior for static variables.

**Description**

Forces the static variable inside the component to be initialized when the component reset signal is asserted. This requires an additional write port to the component memory implemented and can increase the power-up latency when the component is reset.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/component_memories/static_var_init.

**hls_init_on_powerup** Memory Attribute

**Syntax**

hls_init_on_powerup

**Constraints**

N/A

**Default Value**

N/A

**Description**

Sets the component memory implementing the static variable to initialize on power-up when the FPGA is programmed. When the component is reset, the component memory is not reset back to the initialized value of the static.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/component_memories/static_var_init.

**hls_max_concurrency** Memory Attribute

**Deprecated:** This attribute is deprecated and will be removed in a future release. Use the hls_private_copies memory attribute instead.

**Syntax**

hls_max_concurrency(N)

**Constraints**

N/A

**Default Value**

N/A
**Description**
Specifies that the memory implementing the variable or array has \( N \) private copies to allow \( N \) concurrent iterations of a loop at any given time.

Apply this attribute only when the scope of a variable (through its declaration or access pattern) is limited to a loop. If the loop has the `max_concurrency` pragma applied to it, the number of private copies created is the lesser of the `hls_max_concurrency` memory attribute value and the `max_concurrency` pragma value.

**hls_max_replicates Memory Attribute**

**Syntax**
`hls_max_replicates(N)`

**Constraints**
N/A

**Default Value**
N/A

**Description**
Specifies that the memory implementing the variable or array has no more than the \( N \) replicates to enable simultaneous reads from the datapath.

To learn more, review the following tutorial:
`<quartus_installdir>/hls/tutorials/component_memories/memory_bank_configuration`.

**hls_private_copies Memory Attribute**

**Syntax**
`hls_private_copies(N)`

**Constraints**
N/A

**Default Value**
N/A

**Description**
Specifies that the memory implementing the variable or array has \( N \) private copies to allow \( N \) concurrent iterations of a loop at any given time.

Apply this attribute only when the scope of a variable (through its declaration or access pattern) is limited to a loop. If the loop has the `max_concurrency` pragma applied to it, the number of private copies created is the lesser of the `hls_private_copies` memory attribute value and the `max_concurrency` pragma value.
13.7. Intel HLS Compiler Pro Edition Loop Pragmas

Use the Intel HLS Compiler loop pragmas to control how the compiler pipelines the loops in your component.

**Table 44. Intel HLS Compiler Pro Edition Loop Pragmas Summary**

<table>
<thead>
<tr>
<th>Pragma</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>disable_loop_pipelining</td>
<td>Prevents compiler from pipelining a loop,</td>
</tr>
<tr>
<td>ii</td>
<td>Forces a loop to have a loop initiation interval (II) of a specified value.</td>
</tr>
<tr>
<td>ivdep</td>
<td>Ignores memory dependencies between iterations of this loop.</td>
</tr>
<tr>
<td>loop_coalesce</td>
<td>Tries to fuse all loops nested within this loop into a single loop.</td>
</tr>
<tr>
<td>loop_fuse</td>
<td>Directs the compiler to try and fuse pairs of adjacent loops.</td>
</tr>
<tr>
<td>max_concurrency</td>
<td>Limits the number of iterations of a loop that can simultaneously execute at any time.</td>
</tr>
<tr>
<td>max_interleaving</td>
<td>Controls whether iterations of a pipelined inner loop in a loop nest from one invocation of the inner loop can be interleaved in the component data pipeline with iterations from other invocations of the inner loop.</td>
</tr>
<tr>
<td>nofusion</td>
<td>Prevents the annotated loop from being fused with adjacent loops.</td>
</tr>
<tr>
<td>speculated_iterations</td>
<td>Specifies the number of clock cycles that a loop exit condition can take to compute.</td>
</tr>
<tr>
<td>unroll</td>
<td>Unrolls the loop completely or by a number of times.</td>
</tr>
</tbody>
</table>

**disable_loop_pipelining Loop Pragma**

**Syntax**

```
#pragma disable_loop_pipelining
```

**Description**

Tells the compiler to not pipeline this loop.

Disable loop pipelining for a loop when the loop-carried dependencies cause the loop iterations to effectively execute sequentially. With loop pipelining disabled, the Intel HLS Compiler can generate a simpler datapath and reduce the FPGA area utilization of your component.

**Example:**

```c
#pragma disable_loop_pipelining
for (int i = 1; i < N; i++) {  int j = a[i-1];
  // Memory dependency induces a high-latency loop feedback path
  a[i] = foo(j)
}
```

**ii Loop Pragma**

**Syntax**

```
#pragma ii N
```

**Description**

Forces the loop to which you apply this pragma to have a loop initiation interval (II) of \(<N>\), where \(<N>\) is a positive integer value.
Forcing a loop II value can have an adverse effect on the $f_{\text{MAX}}$ of your component because using this pragma to get a lower loop II combines pipeline stages together and creates logic with a long propagation delay.

Using this pragma with a larger loop II inserts more pipeline stages and can give you a better component $f_{\text{MAX}}$ value.

Example:

```c
#pragma ii 2
for (int i = 0; i < 8; i++) {
    // Loop body
}
```

**ivdep Loop Pragma**

**Syntax**

```
#pragma ivdep safelen(N) array(array_name)
```

**Description**

Tells the compiler to ignore memory dependencies between iterations of this loop.

It can accept an optional argument that specifies the name of the array. If `array` is not specified, all component memory dependencies are ignored. If there are loop-carried dependencies, your generated RTL produces incorrect results.

The `safelen` parameter specifies the dependency distance. The dependency distance is the number of iterations between successive load/stores that depend on each other. It is safe to not include `safelen` is only when the dependence distance is infinite (that is, there are no real dependencies).

Example:

```c
#pragma ivdep safelen(2)
for (int i = 0; i < 8; i++) {
    // Loop body
}
```

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/best_practices/loop_memory_dependency`.

**loop_coalesce Loop Pragma**

**Syntax**

```
#pragma loop_coalesce N
```

**Description**

Tells the compiler to try to fuse all loops nested within this loop into a single loop. This pragma accepts an optional value $N$ which indicates the number of levels of loops to coalesce together.

Example:

```c
#pragma loop_coalesce 2
for (int i = 0; i < 8; i++) {
    for (int j = 0; j < 8; j++) {
```
loop_fuse Block-Scope LoopPragma

**Syntax**

```
#pragma loop_fuse [depth(N)] [independent]
```

**Description**

Apply this pragma to a block of code to indicate to the compiler that adjacent loops in the code block should be fused when safe, overriding the compiler profitability analysis of the fusion.

The `depth(N)` clause sets the number of nesting depths the compiler should consider when fusing adjacent loops. Specifying `depth(1)` is equivalent to indicating that only adjacent top-level loops should be considered for fusing.

The `independent` clause overrides the safety checks. If you specify the `independent` option, you are guaranteeing to the compiler that fusing pairs of loops affected by the `loop_fuse` pragma is safe. If it is not safe, you might get functional errors in your component.

For details of the safety checks, see the Fusion Criteria section of Loop Fusion on page 71.

**Example:**

```c
#pragma loop_fuse
{
  for (int j=0; j < N; ++j){
    data[j] += Q;
  }
  for (int i = 0; i < N; ++i){
    output[i] = Q * data[i];
  }
}
```

max_concurrency LoopPragma

**Syntax**

```
#pragma max_concurrency N
```

**Description**

This pragma limits the number of iterations of a loop that can simultaneously execute at any time.

This pragma is useful mainly when private copies are created to improve the throughput of the loop. This is mentioned in the details pane for the loop in the Loop Analysis pane and the Bank view of the Function Memory Viewer of the high level design report (report.html).

This can occur only when the scope of a component memory (through its declaration or access pattern) is limited to this loop. Adding this pragma can be used to reduce the area that the loop consumes at the cost of some throughput.
Example:

```plaintext
// Without this pragma,
// multiple private copies
// of the array "arr"
#pragma max_concurrency 1
for (int i = 0; i < 8; i++) {
    int arr[1024];
    // Loop body
}
```

**max_interleaving Loop Pragma**

**Syntax**

```plaintext
#pragma max_interleaving <option>
```

**Description**

This pragma controls whether iterations of a pipelined inner loop in a loop nest from one invocation of the inner loop can be interleaved in the component data pipeline with iterations from other invocations of the inner loop.

By default, the Intel HLS Compiler tries interleave a number simultaneous invocations of the inner loop equal to the loop initiation interval (II) of the inner loop. For example, an inner loop with an II of 2 can have iterations from two invocations in the pipeline at a time.

In cases where the interleaving of loop iterations from different loop invocations does not yield a performance benefit, limiting or restricting the amount of interleaving can result in reduced FPGA area utilization.

**Supported values for <option>:**

- **1**
  
  The compiler restricts the annotated (inner) loop to be invoked only once per outer loop iteration. That is, all iterations of the inner loop travel the pipeline before the next invocation of the inner loop can occur.

- **0**
  
  Use the default interleaving behavior.

**Example:**

```plaintext
// Loop j is pipelined with ii=1
for (int j = 0; j < M; j++) {
    int a[M];
    // Loop i is pipelined with ii=2
    #pragma max_interleaving 1
    for (int i = 1; i < N; i++) {
        a[i] = foo(i)
    }
    ...
}
```
**nofusion Loop Pragma**

**Syntax**

```c
#pragma nofusion
```

**Description**

This pragma directs the compiler to not fuse the annotated loop with any adjacent loops.

**Example:**

```c
#pragma nofusion
L1: for (int j=0; j < N; ++j){
data[j] += Q;
}
L2: for (int i = 0; i < N; ++l) {
output[i] = Q * data[i];
}
```

**speculated_iterations Loop Pragma**

**Syntax**

```c
#pragma speculated_iterations N
```

**Description**

This pragma specifies the number of loop iterations to wait before considering a loop exit condition. That is, you estimate that a loop takes at least N loop iterations before the exit condition is met.

If you specify a value that is too low, then the loop II increases to accommodate the iterations required to determine whether the loop exit condition is met.

**Example:**

```c
cOMPONENT int loop_speculate (int N) {
    int m = 0;
    // The exit path has 2 multiplies and
    // compare is most critical in loop feedback path
    #pragma speculated_iterations 2
    while (m*m*m < N) {
        m += 1;
    }
    return m;
}
```

**unroll Loop Pragma**

**Syntax**

```c
#pragma unroll N
```

**Description**

This pragma unrolls the loop completely or by <N> times, where <N> is optional and is a positive integer value.

**Important:** Unrolling nested loops with large bounds might generate a large number of instructions that could result in very long compile times for your component.
Example:

```c
#pragma unroll 8
for (int i = 0; i < 8; i++) {
   // Loop body
}
```

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/best_practices/resource_sharing_filter`.

### 13.8. Intel HLS Compiler Pro Edition Scope Pragmas

Use the Intel HLS Compiler scope pragmas to influence the rounding of floating-point operations and the ordering of arithmetic operations in your component at finer grain than the `i++` command options.

<table>
<thead>
<tr>
<th>Pragma</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>fp contract</code></td>
<td>Controls the removal of intermediate rounding and conversion when possible within the code block that this pragma is applied to.</td>
</tr>
<tr>
<td><code>fp reassoc</code></td>
<td>Deprecated: This pragma is deprecated and will be removed in a later release of the Intel HLS Compiler. Use the <code>fp reassociate scoped</code> pragma instead. Controls the relaxing of the order of floating point arithmetic operations within the code block that this pragma is applied to.</td>
</tr>
<tr>
<td><code>fp reassociate</code></td>
<td>Controls the relaxing of the order of floating point arithmetic operations within the code block that this pragma is applied to.</td>
</tr>
</tbody>
</table>

**fp contract Scoped Pragma**

**Syntax**

```c
#pragma clang fp contract(state)
```

**Description**

This pragma controls whether the compiler can contract floating-point multiply and add or subtract operations into a single fused multiply-add (FMA), and controls whether the compiler skips intermediate rounding and conversions.

If multiple occurrences of this pragma affect the same scope, the pragma with the narrowest scope takes precedence.

The `state` parameter can be one of the following values:
• off
  Turns off any permissions to fuse instructions into FMAs.
  The effect of the `-ffp-contract=fast i++` command flag is suppressed for instructions within the scope of the pragma.
• fast
  Allows the fusing of `mult` and `add` instructions into an FMA, but might violate the language standard.
  For instructions with the scope of this pragma, the same optimizations as `-ffp-contract=fast i++` command flag are enabled.

`fp reassoc` ScopedPragma

*Deprecated:* This pragma is deprecated and will be removed in a later release of the Intel HLS Compiler. Use the `fp reassociate scoped` pragma instead.

**Syntax**
```
#pragma clang fp reassoc(state)
```

**Description**
This pragma controls whether the compiler can relax the order of floating point operations requested by the source code. With some reordering, the compiler can optimize the hardware structure which improves the performance of your component.

If multiple occurrences of this pragma affect the same scope, the pragma with the narrowest scope takes precedence.

The `state` parameter can be one of the following values:
• on
  Enables the effect of the `-ffp-reassoc i++` command flag for instructions within the scope of the pragma.
• off
  The effect of the `-ffp-reassoc i++` command flag is suppressed for instructions within the scope of the pragma.

`fp reassociate` ScopedPragma

**Syntax**
```
#pragma clang fp reassociate(state)
```

**Description**
This pragma controls whether the compiler can relax the order of floating point operations requested by the source code. With some reordering, the compiler can optimize the hardware structure which improves the performance of your component.

If multiple occurrences of this pragma affect the same scope, the pragma with the narrowest scope takes precedence.

The `state` parameter can be one of the following values:
on
Enables the effect of the -ffp-reassociate i++ command flag for instructions within the scope of the pragma.

off
The effect of the -ffp-reassociate i++ command flag is suppressed for instructions within the scope of the pragma.

13.9. Intel HLS Compiler Pro Edition Component Attributes

Table 46. Intel HLS Compiler Component Attributes Summary

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hls_component_ii</td>
<td>Force the component to which you apply this attribute to have a specified component initiation interval (II).</td>
</tr>
<tr>
<td>hls_disable_component_pipelining</td>
<td>Prevents the creation of the pipelined component datapath. Multiple invocations of this component now occur sequentially and not simultaneously.</td>
</tr>
<tr>
<td>hls_max_concurrency</td>
<td>Request more copies of the component memory so that the component can run multiple invocations in parallel.</td>
</tr>
<tr>
<td>hls_scheduler_target_fmax_mhz</td>
<td>Specify the target clock frequency of your component.</td>
</tr>
<tr>
<td>hls_use_stall_enable_clusters</td>
<td>Group related operations into stall-enabled clusters to try to improve latency and area usage while possibly sacrificing throughput, when compared to the default stall-free clustering implementation.</td>
</tr>
</tbody>
</table>

### hls_component_ii Component Attribute

**Syntax**

hls_component_ii(<N>)

**Description**

Forces the component to which you apply this attribute to have a component initiation interval (II) of <N>, where <N> is a positive integer value.

This can have an adverse effect on the f MAX of your component because using this attribute to get a lower II combines pipeline stages together and creates logic with a long propagation delay.

Using this attribute with a larger II inserts more pipeline stages and can give you a better component f MAX value.

### hls_disable_component_pipelining Component Attribute

**Syntax**

hls_disable_component_pipelining

**Description**

Tells the compiler to not create a pipelined datapath for the component. An unpipelined component datapath can save FPGA area utilization in some cases.
Use this attribute when a pipelined datapath does not improve your component throughput or when the component is not invoked repeatedly.

Example

```c
#include "HLS/hls.h"

hls_disable_component_pipelining

component void baz ( /* arguments */ ){
    // component code
}
```

### hls_max_concurrency Component Attribute

**Syntax**

```
void hls_max_concurrency(<N>)
```

**Description**

In some cases, the concurrency of a component is limited to 1. This limit occurs when the generated hardware cannot be shared across component invocations. For example, when using component memories for a non-static variable.

You can use this attribute to request more copies of the component memory so that the component can run multiple invocations in parallel.

This attribute can accept any non-negative whole number, including 0.

**Value greater than 0**

A value greater than 0 indicates how many copies of the component memory to instantiate as well as how many component invocations can be in flight at once.

**Value equal to 0**

Setting `hls_max_concurrency` to a value of 0 is useful in cases when there is no component memory but the component still has a poor dynamic loop initiation interval (II) even if you believe your component II should be 1. You can review the II for loops in your component in the high level design report.

To learn more, review the design example: `<quartus_installdir>/hls/examples/inter_decim_filter`.

**Example**

```
hls_max_concurrency(2)

component void foo(ihc::stream_in<int> &data_in,
    ihc::stream_out<int> &data_out) {
    int arr[N];
    for (int i = 0; i < N; i++) {
        arr[i] = data_in.read();
    }
    // Operate on the data and modify in place
    for (int i = 0; i < N; i++) {
        data_out.write(arr[i]);
    }
}
```
**hls_scheduler_target_fmax_mhz Component Attribute**

*Syntax*

\[ \text{hls_scheduler_target_fmax_mhz}(\langle N \rangle) \]

*Description*

Apply the \texttt{hls_scheduler_target_fmax_mhz} component attribute to have the compiler target a specific f\textsubscript{MAX} value. Specify the target f\textsubscript{MAX} value in MHz.

The component is not guaranteed to close timing at the specified frequency, and any tasks in a system of tasks use the same clock regardless of having different scheduling targets.

**hls_use_stall_enable_clusters Component Attribute**

*Syntax*

\[ \text{hls_use_stall_enable_clusters} \]

*Description*

Apply the \texttt{hls_use_stall_enable} component attribute to reduce the area of your component while possibly decreasing your component f\textsubscript{MAX} and throughput.

The Intel HLS Compiler typically groups related operations into clusters. In many cases, the clusters are stall-free clusters. A stall-free cluster executes the operations without any stalls and contains a FIFO at the end of the cluster that holds the results if the cluster is stalled. This FIFO adds area and latency to the component, but might allow a higher f\textsubscript{MAX} and increased throughput.

If you prefer lower FPGA area usage and lower latency over higher throughput, use the \texttt{hls_use_stall_enable} component attribute to bias the compiler to produce stall-enabled clusters. Stall-enabled clusters lack the FIFO, which reduces area and latency, but pass stall signals to the contained operations.

Passing stall signals might reduce f\textsubscript{MAX}.

Not all operations support stall, and these operations cannot be contained in a stall-enabled cluster. The compiler generates a warning if some operations cannot be placed into a stall-enabled cluster.

The compiler automatically uses stall-enabled clusters for HLS components if it can determine that stall-enable is always beneficial. This attribute requests the compiler to form stall-enabled clusters if possible.

*Intel Agilex and Stratix 10 Restriction:*

This attribute does not apply to designs that target Intel Agilex or Intel Stratix 10 devices unless you specify the \texttt{--hyper-optimized-handshaking=off} option of the Intel HLS Compiler \texttt{i++} command.

To learn more, review the following tutorial:

\[ <\text{quartus_installdir}>/hls/examples/tutorials/best_practices/stall_enable \]
13.10. Intel HLS Compiler Pro Edition Component Default Interfaces

Table 47. Intel HLS Compiler Default Interfaces

<table>
<thead>
<tr>
<th>Interface</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Component invocation interface (component call and return)</td>
<td>The component call is implemented as an interface consisting of the component start and busy conduits. The component return is also implemented as an interface that includes the component done and stall signals.</td>
</tr>
<tr>
<td>Scalar parameter interface (passed by value)</td>
<td>Scalar parameters are implemented as input conduits that are synchronized with the component invocation interface.</td>
</tr>
<tr>
<td>Pointer parameter interface (passed by reference)</td>
<td>Pointer parameters are implemented as an implicit Avalon Memory-Mapped Host (mm_host) interface with the default parametrization. By default, the base address is treated as a scalar parameter so it is implemented as a conduit that is synchronized to the component invocation interface. A memory mapped interface is also exposed on the component.</td>
</tr>
</tbody>
</table>

13.11. Intel HLS Compiler Pro Edition Component Invocation Interface Control Attributes

Table 48. Intel HLS Compiler Component Invocation Interface Control Attribute Summary

<table>
<thead>
<tr>
<th>Control Attribute</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hls_avalon_streaming_component</td>
<td>This is the default component invocation interface. The component uses start, busy, stall, and done signals for handshaking.</td>
</tr>
<tr>
<td>hls_avalon_agent_component</td>
<td>The start, done, and returndata (if applicable) signals appear in the component CSR instead of as conduits outside of the signal.</td>
</tr>
<tr>
<td>hls_always_run_component</td>
<td>The start signal is tied to 1 internally in the component. There is no done signal output.</td>
</tr>
<tr>
<td>hls_stall_free_return</td>
<td>The stall signal is removed by internally setting it to 0. Use this control attribute only if the downstream component never stalls.</td>
</tr>
</tbody>
</table>

**hls_avalon_streaming_component Invocation Control Attribute**

**Description**

This is the default component invocation interface. This attribute follows the Avalon ST protocol for both the function call and the return streams. The component consumes the unstable arguments when the start signal is asserted and the busy signal is deasserted. The component produces the return data when the done signal is asserted.
Top-Level Module Ports

- Function call:
  - start
  - busy
- Function return:
  - done
  - stall

Example

```c
component hls_avalon_streaming_component void foo(/*component parameters*/)
```

hls_avalon_agent_component Invocation Control Attribute

Description

The start, done, and return data (if applicable) signals are registered in the component agent memory map. Because the signals are registered in the memory map, the start/busy and stall/done handshaking signals are also removed. The removal of these handshaking signals also removes the handshaking signal for the input parameters.

For signals to be synchronized properly, each of the component parameters must be one of the following parameter types:

- Agent register argument (hls_avalon_agent_register_argument), so that the signals are in the register map. This includes Avalon MM host or pointer interfaces that have the hls_avalon_agent_register_argument parameter applied.

- Agent memory argument (hls_avalon_agent_memory_argument) so that a dedicated Avalon MM agent interface for handshaking is created.

- Stable argument (hls_stable_argument), to explicitly indicate that the signals do not require handshaking. This includes Avalon MM host and pointer interfaces that have the hls_stable_argument parameter applied.

- Streaming interface arguments, so that a dedicated Avalon ST interface for handshaking is created.

If you do not specify one of these component parameters, the compiler generates an error message when you compile this component.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/interfaces/mm_agents`. 

Top-Level Module Ports

- Avalon MM agent interface
- irq_done signal
**Example**

```c
component hls_avalon_agent_component void foo(/*component parameters*/)
```

**hls_always_run_component Invocation Control Attribute**

**Description**

The `start` signal is tied to 1 internally in the component. There is no `done` signal output. The control logic is optimized away when Intel Quartus Prime compiles the generated RTL for your FPGA.

Use this protocol when the component data path relies only on explicit streams for data input and output.

IP verification does not support components with this component invocation protocol.

**Top-Level Module Ports**

None

**Example**

```c
component hls_always_run_component void foo(/*component parameters*/)
```

**hls_stall_free_return Invocation Control Attribute**

**Description**

If the downstream component never stalls, the `stall` signal is removed by internally setting it to 0.

This feature can be used with the `hls_avalon_streaming_component`, `hls_avalon_agent_component`, and `hls_always_run_component` arguments. This attribute can be used to specify that the downstream component is stall free.

**Top-Level Module Ports**

N/A

**Example**

```c
component hls_stall_free_return int dut(int a, int b)
{ return a * b; }
```

---


**Table 49. Intel HLS Compiler Component Macros Summary**

<table>
<thead>
<tr>
<th>Macro</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>hls_avalon_agent_register_argument</code></td>
<td>Implement the parameter as a register that can be read from and written to over an Avalon memory-mapped (MM) agent interface.</td>
</tr>
<tr>
<td><code>hls_avalon_agent_memory_argument</code></td>
<td>Implement the parameter, in on-chip memory blocks, which can be read from or written to over a dedicated agent interface.</td>
</tr>
</tbody>
</table>

*continued*...
<table>
<thead>
<tr>
<th>Macro</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>hls_conduit_argument</td>
<td>Implement the parameter as an input conduit that is synchronous to the component call (start and busy).</td>
</tr>
<tr>
<td>hls_readwrite_mode</td>
<td>Indicate to the compiler how the agent memory interface is accessed by external Avalon memory-mapped (MM) hosts.</td>
</tr>
<tr>
<td>hls_stable_argument</td>
<td>A stable parameter is a parameter that does not change while there is live data in the component (that is, the argument does not change between pipelined function invocations).</td>
</tr>
</tbody>
</table>

**hls_avalon_agent_register_argument Component Macro**

**Syntax**

hls_avalon_agent_register_argument

**Description**

The compiler implements the parameter as a register that can be read from and written to over an Avalon MM agent interface. The parameter will be read into the component pipeline, similar to the conduit implementation. The implementation is synchronous to the start and busy interface.

Changes to the value of this parameter made by the component data path will not be reflected on this register.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/interfaces/mm_agents`.

**Example**

```c
component void foo(
    hls_avalon_agent_register_argument int b)
```

**hls_avalon_agent_memory_argument Component Macro**

**Syntax**

hls_avalon_agent_memory_argument(N)

**Description**

The compiler implements the parameter, where N specifies the size of the memory in bytes, in on-chip memory blocks, which can be read from or written to over a dedicated agent interface. The generated memory has the same architectural optimizations as all other internal component memories (such as banking or coalescing).

If the compiler performs static coalescing optimizations, the agent interface data width is the coalesced width. This attribute applies only to a pointer parameter.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/interfaces/mm_agents`.

**Example**

```c
component void foo(
    hls_avalon_agent_memory_argument(128*sizeof(int)) int *a)
```
**hls_conduit_argument Component Macro**

**Syntax**

hls_conduit_argument

**Description**

This is the default interface for scalar parameters. The compiler implements the parameter as an input conduit that is synchronous to the component call (start and busy).

**Example**

```c
component void foo(hls_conduit_argument int b)
```

**hls_readwrite_mode Component Macro**

**Syntax**

hls_readwrite_mode("type")

**Description**

This macro applies only to agent memory interfaces. Indicates to the compiler how the agent memory interface is accessed by external memory hosts. This information can help the compiler build a more efficient memory system and might save FPGA area for your component.

The `type` parameter can take any one of the following values:

- **readonly**
  Indicates that the external Avalon memory-mapped (MM) host interface only ever reads from the agent memory.

- **writeonly**
  Indicates that the external Avalon MM host interface only ever writes to the agent memory.

**Example**

```c
component void foo(hls_avalon_agent_memory_argument(128*sizeof(int))
    hls_readwrite_mode("writeonly") int *A)
```

**hls_stable_argument Component Macro**

**Syntax**

hls_stable_argument

**Description**

A stable parameter is a parameter that does not change while there is live data in the component (that is, the component argument does not change between pipelined function invocations).

Changing a stable parameter during component execution results in undefined behavior; each use of the stable parameter might be the old value or the new value, but with no guarantee of consistency. The same variable in the same invocation can appear with multiple values.

Using stable parameters, where appropriate, might save a significant number of registers in a design.
Stable parameters can be used with conduits, Avalon MM host interfaces, and agent_registers.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/interfaces/stable_arguments`.

**Example**

```
component int dut(
  hls_stable_argument int a,
  hls_stable_argument int b) {
  return a * b;
}
```

# 13.13. Intel HLS Compiler Pro Edition Systems of Tasks API

## Table 50. Intel HLS Compiler System of Tasks Summary

<table>
<thead>
<tr>
<th>Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ihc::launch</code></td>
<td>Marks a function as an Intel HLS Compiler task for hardware generation, and launches the task function asynchronously.</td>
</tr>
<tr>
<td><code>ihc::collect</code></td>
<td>Synchronizes the completion of the specified task function in the component.</td>
</tr>
<tr>
<td><code>ihc::stream</code></td>
<td>Allows streaming communication between different task functions.</td>
</tr>
<tr>
<td><code>ihc::launch_always_run</code></td>
<td>Launches a task function at component power-on or reset and continuously executes the function.</td>
</tr>
<tr>
<td></td>
<td><strong>Recommendation:</strong> Use the <code>ihc_hls_set_component_wait_cycle</code> with this function to keep your component and always-run task functions correctly coordinated.</td>
</tr>
</tbody>
</table>

### `ihc::launch` Function

**Syntax**

```
ihc::launch<function[, capacity]>(function_argument_list)
```

Where the function parameters are defined as follows:

- **function**
  The name of the function that you are calling as an Intel HLS Compiler task in your component.

- **capacity**
  An optional value that, when set, results in a FIFO buffer of depth `capacity` inserted between the function that launches the task and the task function.

  Set the `capacity` parameter when you observe stall patterns that indicate an imbalance between any backpressure introduced by the called task function (`function`) and how often the caller launches this task function.

- **function_argument_list**
  The list of arguments to pass to the task function.

  This list must match the arguments (in names and types) that the task function expects.
The `ihc::launch` API function identifies a function as Intel HLS Compiler task for hardware generation. Calling this function starts the task function asynchronously.

If the task function cannot accept a new thread, the `ihc::launch` function can block the function that calls the `ihc::launch` function.

The list of arguments that supply the `ihc::launch` API function must match (in names and types) the list of arguments expected by the task function.

### `ihc::collect` Function

#### Syntax

```
ihc::collect<function[, capacity]>()
```

Where the function parameters are defined as follows:

- **function**
  The name of the Intel HLS Compiler task function to synchronize the completion of.

  Set the `capacity` parameter when you observe stall patterns that indicate that the task function (`function`) produces data at a different cadence from the reading cadence of the caller.

- **capacity**
  An optional value that, when set, results in a FIFO buffer of size `capacity` inserted between the task function and the function that collects the task.

#### Description

The `ihc::collect` API function synchronizes the completion of the specified task function in the component.

For a non-void task function, the `ihc::collect` API function collects the result from the specified task function.

For a void task function, the `ihc::collect` API function synchronizes against the `done` signal of the task function.

The number of `ihc::collect` calls for a task function must match the number of `ihc::launch` calls for the same task function to flush all of the calls to the task.

**Special Case:** If you do not use `ihc::collect` at all, the compiler optimizes and ties-off the return stream of the task to be stall free and ignores any data on the return stream. Other streaming interfaces can still back-pressure the task function. Additionally, the caller might finish before the task function.

### `ihc::launch_always_run` Function

#### Syntax

```
ihc::launch_always_run<function>()
```

Where the function parameters are defined as follows:

- `<function>`
  
  The name of the function that you are calling as a continuously-executing Intel HLS Compiler task in your component.

**Description**

Use the `ihc::launch_always_run` API function to continuously execute a task function, much like an invoking a component with the `hls_always_run_component` invocation interface argument.

The task launches at the power-on or the reset of the component instead of at a specific point in the datapath.

The task function that you provide to this API must match this prototype:

```c
void function(void)
```

Your task function must be have no function arguments and no return value. You should communicate with your task function through global streams or by using compile-time constant template parameters.

**Example**

The following example shows a simple use of the `ihc::launch_always_run` function.

```c
ihc::stream<int> in_stream, out_stream;

template <ihc::stream<int> &inStream, 
          ihc::stream<int> &outStream>
void my_task()
{
    int x = inStream.read();
    x *= 2;
    outStream.write(x);
}

component void foo()
{
    ihc::launch_always_run<my_task<in_stream, out_stream>>();
}
```

**Intel HLS Compiler System of Tasks Code Example**

The following code example illustrates how you can use the systems of tasks API.

```c
int mul(int a, int b)
{
    return a * b;
}

template<typename T>
T add(T a, T b)
{
    return a + b;
}
```
component int foo(int a, int b)
{
    ihc::launch<mul>(a,b);
    ihc::launch<add<int>>(a,b);
    int prod = ihc::collect<mul>();
    int sum = ihc::collect<add<int>>();
    return sum + prod;
}

Related Information
Intel HLS Compiler Pro Edition Component Invocation Interface Control Attributes on page 171

13.13.1. ihc::stream Class

Table 51. Intel HLS Compiler Systems of Tasks Streaming Interface Template Summary

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::stream</td>
<td>Streaming interface to the component or task function.</td>
</tr>
<tr>
<td>ihc::buffer</td>
<td>Specifies the capacity (in words) of the FIFO buffer on the input data that associates with the stream.</td>
</tr>
<tr>
<td>ihc::usesPackets</td>
<td>Exposes the startofpacket and endofpacket sideband signals on the stream interface.</td>
</tr>
</tbody>
</table>

**ihc::stream Template Object**

*Syntax*  
ihc::stream<datatype, template arguments>

*Valid Values*  
Any trivially copyable C++ data type.

*Default Value*  
N/A

*Description*  
Streaming interface to the component or task. The width of the stream data bus is equal to a width of sizeof(datatype).

**ihc::buffer Template Parameter**

*Syntax*  
ihc::buffer<value>

*Valid Values*  
Non-negative integer value.

*Default Value*  
0

*Description*  
The capacity, in words, of the FIFO buffer on the input data that associates with the stream.
**ihc::usesPackets Template Parameter**

**Syntax**

ihc::usesPackets<value>

**Valid Values**

true or false

**Default Value**

false

**Description**

Exposes the startofpacket and endofpacket sideband signals on the stream interface, which can be accessed by the packet based reads/writes.

**Intel HLS Compiler System of Tasks Streaming Interface stream Function APIs**

<table>
<thead>
<tr>
<th>Function API</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T read()</td>
<td>Blocking read call to be used from within the component or task</td>
</tr>
<tr>
<td>T read(bool&amp; sop, bool&amp; eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Blocking read with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>T tryRead(bool &amp;success)</td>
<td>Non-blocking read call to be used from within the component or task. The success bool is set to true if the read was valid.</td>
</tr>
<tr>
<td>T tryRead(bool&amp; success, bool&amp; sop, bool&amp; eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Non-blocking read with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>void write(T data)</td>
<td>Blocking write call from the component or task.</td>
</tr>
<tr>
<td>void write(T data, bool sop, bool eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Blocking write with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>bool tryWrite(T data)</td>
<td>Non-blocking write call from the component or task. The return value represents whether the write was successful.</td>
</tr>
<tr>
<td>bool tryWrite(T data, bool sop, bool eop)</td>
<td>Available only if usesPackets&lt;true&gt; is set. Non-blocking write with out-of-band startofpacket and endofpacket signals. The return value represents whether the write was successful.</td>
</tr>
</tbody>
</table>


The pipe API is equivalent to the following class declaration:

```cpp
template <class name, class dataT, size_t min_capacity = 0>
class pipe {
    public:
        // Blocking
        static dataT read();
        static void write(dataT data);
        // Non-blocking
```
static dataT read(bool &success);
static void write(dataT data, bool &success);

Where the template parameters are defined as follows:

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>name</td>
<td>The type that is the basis of a pipe identification. It is typically a user-defined class, in a user namespace. Forward declaration of the type is enough, and the type need not be defined.</td>
</tr>
<tr>
<td>dataT</td>
<td>The data type of the packet contained within a pipe. This is the data type that is read during a successful pipe read() operation, or written during a successful pipe write() operation. The type must have a standard layout and be trivially copyable.</td>
</tr>
<tr>
<td>min_capacity</td>
<td>The minimum number of words (in units of dataT) that the pipe must be able to store without any being read out. The compiler might create a pipe with a larger capacity due to performance considerations.</td>
</tr>
</tbody>
</table>

13.15. Intel HLS Compiler Pro Edition Streaming Input Interfaces

Use the stream_in object and template arguments to explicitly declare Avalon Streaming (ST) input interfaces. You can also use the stream_in Function APIs.

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::stream_in</td>
<td>Streaming input interface to the component.</td>
</tr>
<tr>
<td>ihc::buffer</td>
<td>Specifies the capacity (in words) of the FIFO buffer on the input data that associates with the stream.</td>
</tr>
<tr>
<td>ihc::readyLatency</td>
<td>Specifies the number of cycles between when the ready signal is deasserted and when the input stream can no longer accept new inputs.</td>
</tr>
<tr>
<td>ihc::bitsPerSymbol</td>
<td>Describes how the data is broken into symbols on the data bus.</td>
</tr>
<tr>
<td>ihc::firstSymbolInHighOrderBits</td>
<td>Specifies whether the data symbols in the stream are in big endian order.</td>
</tr>
<tr>
<td>ihc::usesPackets</td>
<td>Exposes the startofpacket and endofpacket sideband signals on the stream interface.</td>
</tr>
<tr>
<td>ihc::usesEmpty</td>
<td>Exposes the empty out-of-band signal on the stream interface.</td>
</tr>
<tr>
<td>ihc::usesValid</td>
<td>Controls whether a valid signal is present on the stream interface.</td>
</tr>
</tbody>
</table>

**ihc::stream_in Template Object**

**Syntax**

```
ihc::stream_in<datatype, template parameters>
```

**Valid Values**

Any valid C++ datatype

**Default Value**

N/A

**Description**

Streaming input interface to the component.
The width of the stream data bus is equal to a width of `sizeof(datatype)`.

The testbench must populate this buffer (stream) fully before the component can start to read from the buffer.

To learn more, review the following tutorials:
- `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_buffer`
- `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_packets_empty`
- `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_packet_ready_valid`
- `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_ready_latency`
- `<quartus_installdir>/hls/examples/tutorials/interfaces/multiple_stream_call_sites`

**ihc::buffer Template Parameter**

*Syntax*  
`ihc::buffer<value>`

*Valid Values*  
Non-negative integer value.

*Default Value*  
0

*Description*  
The capacity, in words, of the FIFO buffer on the input data that associates with the stream. The buffer has latency. It immediately consumes data, but this data is not immediately available to the logic in the component.

If you use the `tryRead()` function to access this stream and the stream read is scheduled within the first cycles of operation, the first (or more) calls to the `tryRead()` function might return `false` in simulation (and therefore in hardware).

Review the function viewer in the System Viewer of the High-Level Design Reports to see when operations are scheduled in your component. If you see this behavior, use the blocking `read()` function to ensure consistency between emulation and simulation.

This parameter is available only on input streams.

**ihc::readyLatency Template Parameter**

*Syntax*  
`ihc::readyLatency<value>`

*Valid Values*  
Non-negative integer value between 0-8.
**Default Value**  
0

**Description**  
The number of cycles between when the `ready` signal is deasserted and when the input stream can no longer accept new inputs.

---

**ihc::bitsPerSymbol Template Parameter**

**Syntax**  
`ihc::bitsPerSymbol<value>`

**Valid Values**  
A positive integer value that evenly divides by the data type size.

**Default Value**  
Datatype size

**Description**  
Describes how the data is broken into symbols on the data bus.

Data is broken down according to how you set the `ihc::firstSymbolInHighOrderBits` declaration. By default, data is broken down in little endian order.

---

**ihc::firstSymbolInHighOrderBits Template Parameter**

**Syntax**  
`ihc::firstSymbolInHighOrderBits<value>`

**Valid Values**  
`true` or `false`

**Default Value**  
`false`

**Description**  
Specifies whether the data symbols in the stream are in big endian order.

---

**ihc::usesPackets Template Parameter**

**Syntax**  
`ihc::usesPackets<value>`

**Valid Values**  
`true` or `false`

**Default Value**  
`false`
**Description**

Exposes the startofpacket and endofpacket sideband signals on the stream interface, which can be accessed by the packet based reads/writes.

**ihc::usesEmpty Template Parameter**

**Syntax**

ihc::usesEmpty<value>

**Valid Values**

true or false

**Default Value**

false

**Description**

Exposes the empty out-of-band signal on the stream interface.

Use this declaration only with streams that read more than one data symbol per clock cycle.

The empty signal indicates the number of symbols on the data bus that do not represent valid data during the final stream read of a packet.

You can control whether the empty symbols are in the low-order bits or high-order bits with the ihc::firstSymbolInHighOrderBits declaration.

**ihc::usesValid Template Parameter**

**Syntax**

ihc::usesValid<value>

**Valid Values**

true or false

**Default Value**

true

**Description**

Controls whether a valid signal is present on the stream interface. If false, the upstream source must provide valid data on every cycle that ready is asserted.

This is equivalent to changing the stream read calls to tryRead and assuming that success is always true.

If set to false, buffer and readyLatency must be 0.
### Intel HLS Compiler Pro Edition Streaming Input Interface `stream_in` Function APIs

#### Table 55. Intel HLS Compiler Pro Edition Streaming Input Interface `stream_in` Function APIs

<table>
<thead>
<tr>
<th>Function API</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T read()</td>
<td>Blocking read call to be used from within the component.</td>
</tr>
<tr>
<td>T read(bool&amp; sop, bool&amp; eop)</td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> is set. Blocking read with out-of-band <code>startofpacket</code> and <code>endofpacket</code> signals.</td>
</tr>
<tr>
<td>T read(bool&amp; sop, bool&amp; eop, int&amp; empty)</td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> and <code>usesEmpty&lt;true&gt;</code> are set. Blocking read with out-of-band <code>startofpacket</code>, <code>endofpacket</code>, and <code>empty</code> signals.</td>
</tr>
<tr>
<td>T tryRead(bool &amp;success)</td>
<td>Non-blocking read call to be used from within the component. The <code>success</code> bool is set to true if the read was valid. That is, the Avalon-ST <code>valid</code> signal was high when the component tried to read from the stream. The emulation model of <code>tryRead()</code> is not cycle-accurate, so the behavior of <code>tryRead()</code> might differ between emulation and simulation.</td>
</tr>
<tr>
<td>T tryRead(bool&amp; success, bool&amp; sop, bool&amp; eop)</td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> is set. Non-blocking read with out-of-band <code>startofpacket</code> and <code>endofpacket</code> signals.</td>
</tr>
<tr>
<td>T tryRead(bool&amp; success, bool&amp; sop, bool&amp; eop, int&amp; empty)</td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> and <code>usesEmpty&lt;true&gt;</code> are set. Non-blocking read with out-of-band <code>startofpacket</code>, <code>endofpacket</code>, and <code>empty</code> signals.</td>
</tr>
<tr>
<td>void write(T data)</td>
<td>Blocking write call to be used from the testbench to populate the FIFO to be sent to the component.</td>
</tr>
<tr>
<td>void write(T data, bool sop, bool eop)</td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> is set. Blocking write call with out-of-band <code>startofpacket</code> and <code>endofpacket</code> signals.</td>
</tr>
<tr>
<td>void write(T data, bool sop, bool eop, int empty)</td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> and <code>usesEmpty&lt;true&gt;</code> are set. Blocking write call with out-of-band <code>startofpacket</code>, <code>endofpacket</code>, and <code>empty</code> signals.</td>
</tr>
</tbody>
</table>

### Intel HLS Compiler Streaming Input Interfaces Code Example

The following code example illustrates both `stream_in` declarations and `stream_in` function APIs.

```c++
// Blocking read
void foo (ihc::stream_in<int> &a) {
    int x = a.read();
}

// Non-blocking read
void foo_nb (ihc::stream_in<int> &a) {
    bool success = false;
    int x = a.tryRead(success);
    if (success) {
        // x is valid
    }
}
```

---


MNL-1083 | 2021.06.21

Send Feedback


185
int main() {
    ihc::stream_in<int> a;
    ihc::stream_in<int> b;
    for (int i = 0; i < 10; i++) {
        a.write(i);
        b.write(i);
    }
    foo(a);
    foo_nb(b);
}

13.16. Intel HLS Compiler Pro Edition Streaming Output Interfaces

Use the stream_out object and template arguments to explicitly declare Avalon Streaming (ST) output interfaces. You can also use the stream_out Function APIs.

Table 56. Intel HLS Compiler Pro Edition Streaming Output Interface Template Summary

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::stream_out</td>
<td>Streaming output interface from the component.</td>
</tr>
<tr>
<td>ihc::readylatency</td>
<td>Specifies the number of cycles between when the ready signal is deasserted and when the input stream can no longer accept new inputs.</td>
</tr>
<tr>
<td>ihc::bitsPerSymbol</td>
<td>Describes how the data is broken into symbols on the data bus.</td>
</tr>
<tr>
<td>ihc::firstSymbolInHighOrderBits</td>
<td>Specifies whether the data symbols in the stream are in big endian order.</td>
</tr>
<tr>
<td>ihc::usesPackets</td>
<td>Exposes the startofpacket and endofpacket sideband signals on the stream interface.</td>
</tr>
<tr>
<td>ihc::usesEmpty</td>
<td>Exposes the empty out-of-band signal on the stream interface.</td>
</tr>
<tr>
<td>ihc::usesReady</td>
<td>Controls whether a ready signal is present.</td>
</tr>
</tbody>
</table>

**ihc::stream_out Template Object**

**Syntax**  
ihc::stream_out<datatype, template parameter>

**Valid Values**  
Any valid POD (plain old data) C++ datatype.

**Default Value**  
N/A

**Description**  
Streaming output interface from the component. The testbench can read from this buffer once the component returns.

To learn more, review the following tutorials:
• `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_buffer`
• `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_packets_empty`
• `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_packet_ready_valid`
• `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_ready_latency`
• `<quartus_installdir>/hls/examples/tutorials/interfaces/multitle_stream_call_sites`

**ihc::readylatency** Template Parameter

**Syntax**

```
ihc::readylatency<value>
```

**Valid Values**

Non-negative integer value (between 0-8)

**Default Value**

0

**Description**

The number of cycles between when the `ready` signal is deasserted and when the sink can no longer accept new inputs.

Conceptually, you can view this parameter as an almost ready latency on the input FIFO buffer for the data that associates with the stream.

**ihc::bitsPerSymbol** Template Parameter

**Syntax**

```
ihc::bitsPerSymbol<value>
```

**Valid Values**

Positive integer value that evenly divides the data type size.

**Default Value**

Datatype size

**Description**

Describes how the data is broken into symbols on the data bus.

Data is broken down according to how you set the `ihc::firstSymbolInHighOrderBits` declaration. By default, data is broken down in little endian order.

**ihc::firstSymbolInHighOrderBits** Template Parameter

**Syntax**

```
ihc::firstSymbolInHighOrderBits<value>
```

**Valid Values**

true or false

**Default Value**

false
**Description**
Specifies whether the data symbols in the stream are in big endian order.

![Big Endian vs Little Endian](image)

### ihc::usesPackets Template Parameter

**Syntax**
```
ihc::usesPackets<value>
```

**Valid Values**
true or false

**Default Value**
false

**Description**
Exposes the start of packet and end of packet sideband signals on the stream interface, which can be accessed by the packet based reads/writes.

### ihc::usesEmpty Template Parameter

**Syntax**
```
ihc::usesEmpty<value>
```

**Valid Values**
true or false

**Default Value**
false

**Description**
Exposes the empty out-of-band signal on the stream interface.

Use this declaration only with streams that write more than one data symbol per clock cycle.

The empty signal indicates the number of symbols on the data bus that do not represent valid data during the final stream write of a packet.

You can control whether the empty symbols are in the low-order bits or high-order bits with the `ihc::firstSymbolInHighOrderBits` declaration.
ihc::usesReady Template Parameter

**Syntax**

```
ihc::usesReady<value>
```

**Valid Values**

true or false

**Default Value**

true

**Description**

Controls whether a ready signal is present. If false, the downstream sink must be able to accept data on every cycle that valid is asserted. This is equivalent to changing the stream read calls to `tryWrite` and assuming that `success` is always true.

If set to false, `readyLatency` must be 0.

Intel HLS Compiler Pro Edition Streaming Output Interface `stream_out` Function APIs

Table 57. Intel HLS Compiler Pro Edition Streaming Output Interface `stream_out` Function APIs

<table>
<thead>
<tr>
<th>Function API</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>void write(T data)</code></td>
<td>Blocking write call from the component</td>
</tr>
<tr>
<td><code>void write(T data, bool sop, bool eop)</code></td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> is set. Blocking write with out-of-band <code>startofpacket</code> and <code>endofpacket</code> signals.</td>
</tr>
<tr>
<td><code>void write(T data, bool sop, bool eop, int empty)</code></td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> and <code>usesEmpty&lt;true&gt;</code> are set. Blocking write with out-of-band <code>startofpacket</code>, <code>endofpacket</code>, and <code>empty</code> signals.</td>
</tr>
<tr>
<td><code>bool tryWrite(T data)</code></td>
<td>Non-blocking write call from the component. The return value represents whether the write was successful.</td>
</tr>
<tr>
<td><code>bool tryWrite(T data, bool sop, bool eop)</code></td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> is set. Non-blocking write with out-of-band <code>startofpacket</code> and <code>endofpacket</code> signals. The return value represents whether the write was successful. That is, the downstream interface was pulling the ready signal high while the HLS component tried to write to the stream.</td>
</tr>
<tr>
<td><code>bool tryWrite(T data, bool sop, bool eop, int empty)</code></td>
<td>Available only if <code>usesPackets&lt;true&gt;</code> and <code>usesEmpty&lt;true&gt;</code> are set. Non-blocking write with out-of-band <code>startofpacket</code>, <code>endofpacket</code>, and <code>empty</code> signals. The return value represents whether the write was successful.</td>
</tr>
</tbody>
</table>
### Function API

<table>
<thead>
<tr>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>T read()</td>
</tr>
<tr>
<td>Blocking read call to be used from the testbench to read back the data from the component.</td>
</tr>
<tr>
<td>T read(bool &amp;sop, bool &amp;eop)</td>
</tr>
<tr>
<td>Available only if usesPackets&lt;true&gt; is set. Blocking read call to be used from the testbench to read back the data from the component with out-of-band startofpacket and endofpacket signals.</td>
</tr>
<tr>
<td>T read(bool &amp;sop, bool &amp;eop, int &amp;empty)</td>
</tr>
<tr>
<td>Available only if usesPackets&lt;true&gt; and usesEmpty&lt;true&gt; are set. Blocking read call to be used from the testbench to read back the data from the component with out-of-band startofpacket, endofpacket, and empty signals.</td>
</tr>
</tbody>
</table>

---

### Intel HLS Compiler Streaming Output Interfaces Code Example

The following code example illustrates both `stream_out` declarations and `stream_out` function APIs.

```cpp
// Blocking write
void foo (ihc::stream_out<int> &a) {
  static int count = 0;
  for(int idx = 0; idx < 5; idx ++){
    a.write(count++);  // Blocking write
  }
}

// Non-blocking write
void foo_nb (ihc::stream_out<int> &a) {
  static int count = 0;
  for(int idx = 0; idx < 5; idx ++){
    bool success = a.tryWrite(count++);  // Non-blocking write
    if (success) {
      // write was successful
    }
  }
}

int main() {
  ihc::stream_out<int> a;
  foo(a); // or foo_nb(a);

  // copy output to an array
  int outputData[5];
  for (int i = 0; i < 5; i++) {
    outputData[i] = a.read();
  }
}
```

---

### 13.17. Intel HLS Compiler Pro Edition Memory-Mapped Interfaces

Use the `mm_host` object and template arguments to explicitly declare Avalon Memory-Mapped (MM) Host interfaces for your component.

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ihc::mm_host</code></td>
<td>The underlying pointer type.</td>
</tr>
<tr>
<td><code>ihc::dwidth</code></td>
<td>The width of the memory-mapped data bus in bits</td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>Template Object or Parameter</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ihc::awidth</td>
<td>The width of the memory-mapped address bus in bits.</td>
</tr>
<tr>
<td>ihc::aspace</td>
<td>The address space of the interface that associates with the host.</td>
</tr>
<tr>
<td>ihc::latency</td>
<td>The guaranteed latency from when a read command exits the component when the external memory returns valid read data.</td>
</tr>
<tr>
<td>ihc::maxburst</td>
<td>The maximum number of data transfers that can associate with a read or write transaction.</td>
</tr>
<tr>
<td>ihc::align</td>
<td>The alignment of the base pointer address in bytes.</td>
</tr>
<tr>
<td>ihc::readwrite_mode</td>
<td>The port direction of the interface.</td>
</tr>
<tr>
<td>ihc::waitrequest</td>
<td>Adds the waitrequest signal that is asserted by the agent when it is unable to respond to a read or write request.</td>
</tr>
<tr>
<td>getInterfaceAtIndex</td>
<td>This testbench function is used to index into an mm_host object.</td>
</tr>
</tbody>
</table>

**ihc::mm_host Template Object**

**Syntax**

ihc::mm_host<datatype, template parameter>

**Valid values**

Any valid C++ datatype

**Default Value**

Default interface for pointer arguments.

**Description**

The underlying pointer type. Pointer arithmetic performed on the host object conforms to this type. Dereferences of the host results in a load-store site with a width of `sizeof(datatype)`. The default alignment is aligned to the size of the datatype.

You can use multiple template arguments in any combination as long the combination of arguments describes a valid hardware configuration.

**Example:**

```cpp
component int dut(
  ihc::mm_host<int,
    ihc::aspace<2>, ihc::latency<3>,
    ihc::awidth<10>, ihc::dwidth<32>
  > &a)
```

To learn more, review the following tutorials:

- `<quartus_installdir>/hls/examples/tutorials/interfaces/pointer_mm_host`
- `<quartus_installdir>/hls/examples/tutorials/interfaces/mm_host_testbench_operators`

**ihc::dwidth Template Parameter**

**Syntax**

ihc::dwidth<value>

**Valid Values**

8, 16, 32, 64, 128, 256, 512, or 1024
Default Value 64

Description The width of the memory-mapped data bus in bits.

**ihc::awidth Template Parameter**

**Syntax** `ihc::awidth<value>`

**Valid Values** Integer value in the range 1 – 64

**Default Value** 64

**Description** The width of the memory-mapped address bus in bits.

This value affects only the width of the Avalon MM Host interface. The size of the conduit of the base address pointer is always set to 64-bits.

**ihc::aspace Template Parameter**

**Syntax** `ihc::aspace<value>`

**Valid Values** Integer value greater than 0.

**Default Value** 1

**Description** The address space of the interface that associates with the host. Each unique `value` results in a separate Avalon MM Host interface on your component. All hosts with the same address space are arbitrated within the component to a single interface. As such, these hosts must share the same template parameters that describe the interface.

**ihc::latency Template Parameter**

**Syntax** `ihc::latency<value>`

**Valid Values** Non-negative integer value

**Default Value** 1

**Description** The guaranteed latency from when a read command exits the component when the external memory returns valid read data. If this latency is variable (such as when accessing DRAM), set it to 0.

**ihc::maxburst Template Parameter**

**Syntax** `ihc::maxburst<value>`

**Valid Values** Integer value in the range 1 – 1024
Default Value 1

Description The maximum number of data transfers that can associate with a read or write transaction. This value controls the width of the burstcount signal.

For fixed latency interfaces, this value must be set to 1.

For more details, review information about burst signals and the burstcount signal role in "Avalon Memory-Mapped Interface Signal Roles" in Avalon Interface Specifications.

**ihc::align Template Parameter**

Syntax ihc::align<value>

Valid Values Integer value greater than the alignment of the datatype

Default Value Alignment of the datatype

Description The alignment of the base pointer address in bytes.

The Intel HLS Compiler uses this information to determine how many simultaneous loads and stores this pointer can permit.

For example, if you have a bus with 4 32-bit integers on it, you should use ihc::dwidth<128> (bits) and ihc::align<16> (bytes). This means that up to 16 contiguous bytes (or 4 32-bit integers) can be loaded or stored as a coalesced memory word per clock cycle.

*Important*: The caller is responsible for aligning the data to the set value for the align argument; otherwise, functional failures might occur.

**ihc::readwrite_mode Template Parameter**

Syntax ihc::readwrite_mode<value>

Valid Values readwrite, readonly, or writeonly

Default Value readwrite

Description The port direction of the interface. Only the relevant Avalon host signals are generated.

**ihc::waitrequest Template Parameter**

Syntax ihc::waitrequest<value>

Valid Values true or false
**Default Value**  
false  

**Description**  
Add the `waitrequest` signal that is asserted by the agent when it is unable to respond to a read or write request. For more information about the `waitrequest` signal, see "Avalon Memory-Mapped Interface Signal Roles" in *Avalon Interface Specifications*.

---

### getInterfaceAtIndex Testbench Function

**Syntax**  
```c
getInterfaceAtIndex(int index)
```

**Description**  
This testbench function is used to index into an `mm_host` object. It can be useful when iterating over an array and invoking a component on different indices of the array. This function is supported only in the testbench.

**Code Example**  
```c
int main() {
    // ……
    for(int idx = 0; idx < N; idx++) {
        dut(src_mm.getInterfaceAtIndex(idx));
    }
    // ……
}
```

---

### 13.18. Intel HLS Compiler Pro Edition Load-Store Unit Control

For variable-latency Avalon Memory-Mapped (MM) Host interfaces (`ihc::latency<0>`), you can control the type of load-store unit (LSU) with the `ihc::lsu` template object and the corresponding `load()` and `store()` functions.

#### Table 59. Intel HLS Compiler Pro Edition Load-Store Unit Control Summary

<table>
<thead>
<tr>
<th>Template Object/Parameter/Function</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>ihc::lsu</code></td>
<td>The underlying LSU class template object</td>
</tr>
<tr>
<td><code>ihc::style</code></td>
<td>Specifies the type of load-store unit.</td>
</tr>
<tr>
<td><code>ihc::static_coalescing</code></td>
<td>Explicitly allows or prevents static coalescing of a load/store operation with other load/store operations.</td>
</tr>
<tr>
<td><code>load</code></td>
<td>Loads data from memory into the LSU.</td>
</tr>
<tr>
<td><code>store</code></td>
<td>Stores data from the LSU into memory.</td>
</tr>
</tbody>
</table>

**ihc::lsu Template Object**

**Syntax**  
`ihc::lsu<template arguments>`

**Valid Values**  
N/A.

**Default Value**  
N/A.
Description  The underlying LSU class object.

To learn more, review the following tutorial:
<quartus_installdir>/hls/examples/tutorials/best_practices/lsu_control

ihc::style Template Parameter

Syntax  ihc::style<LSU_type>

Valid Values  LSU_type can be one of the following values:
- BURST_COALESCED
- PIPELINED

Default Value  BURST_COALESCED

Description  Specifies the type of load-store unit to create.

A burst-coalesced LSU buffers requests until the largest possible burst
can be made.

A pipelined LSU submits requests as they are received.

ihc::static_coalescing Template Parameter

Syntax  ihc::static_coalescing<value>

Valid Values  true or false

Default Value  true

Description  Specifies whether to allow or prevent static coalescing of the load/store operation with other load/store operations.

load Function

Syntax  load(<memory_location>)

Parameters  The <memory_location> argument specifies the memory location to
load data into the LSU from.

Return Type  Object of same type as the base type of the argument specified for
<memory_location>.

Description  The load function loads data from a memory location specified by the
<memory_location> argument and returns the data that the argument
points to.
### store Function

**Syntax**

```c
store(<memory_location>, <value_to_store>)
```

**Parameters**

- The `<memory_location>` argument specifies the memory location to store data coming from the LSU.
- The `<value_to_store>` argument is the value from the LSU to store in memory. The type is the same as the pointer base type.

**Return Type**

None.

**Description**

The `store` function stores data in the LSU to a memory location specified by the `<memory_location>` argument.

### 13.19. Intel HLS Compiler Pro Edition Arbitrary Precision Data Types

**Table 60. Arbitrary Precision Data Types Supported by the Intel HLS Compiler Pro Edition**

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Intel Header File</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ac_int</td>
<td>HLS/ac_int.h</td>
<td>Arbitrary-width integer support. To learn more, review the following tutorials:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>&lt;quartus_installdir&gt;/hls/examples/tutorials/ac_datatypes/ac_int_basic_ops</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>&lt;quartus_installdir&gt;/hls/examples/tutorials/ac_datatypes/ac_int_overflow</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>&lt;quartus_installdir&gt;/hls/examples/tutorials/best_practices/struct_interfaces</code></td>
</tr>
<tr>
<td>ac_fixed</td>
<td>HLS/ac_fixed.h</td>
<td>Arbitrary-precision fixed-point number support. To learn more, review the tutorial: <code>&lt;quartus_installdir&gt;/hls/examples/tutorials/ac_datatypes/ac_fixed_constructor</code></td>
</tr>
<tr>
<td>ac_complex</td>
<td>HLS/ac_complex.h</td>
<td>Complex number support.</td>
</tr>
<tr>
<td>hls_float</td>
<td>HLS/hls_float.h</td>
<td>Arbitrary-precision floating-point number support.</td>
</tr>
<tr>
<td></td>
<td>HLS/hls_float_math.h</td>
<td>Support for commonly used exponential, logarithmic, power, and trigonometric functions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>To learn more, review the following tutorials:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>&lt;quartus_installdir&gt;/hls/examples/tutorials/hls_float/1_reduced_double</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>&lt;quartus_installdir&gt;/hls/examples/tutorials/hls_float/2_explicit_arithmetic</code></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <code>&lt;quartus_installdir&gt;/hls/examples/tutorials/hls_float/3_conversions</code></td>
</tr>
</tbody>
</table>
Table 61. Intel HLS Compiler ac_int Debugging Tools Summary

<table>
<thead>
<tr>
<th>Tool</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEBUG_AC_INT_WARNING</td>
<td>Emits a warning for each detected overflow.</td>
</tr>
<tr>
<td>DEBUG_AC_INT_ERROR</td>
<td>Emits a message for the first overflow that is detected and then exits the component with an error.</td>
</tr>
</tbody>
</table>

**DEBUG_AC_INT_WARNING ac_int Debugging Tool**

**Macro Syntax**

```c
#define DEBUG_AC_INT_WARNING
```

If you use this macro, declare it in your code before you declare `#include HLS/ac_int.h`.

**i++ Command Option Syntax**

```bash
-D DEBUG_AC_INT_WARNING
```

**Description**

Enables runtime tracking of ac_int data types during x86 emulation (the `-march=x86-64` option, which is the default option, of the i++ command).

This tool uses additional resources for tracking the overflow and empty constructors, and emits a warning for each detected overflow.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/ac_datatypes/ac_int_overflow`

**DEBUG_AC_INT_ERROR ac_int Debugging Tool**

**Macro Syntax**

```c
#define DEBUG_AC_INT_ERROR
```

If you use this macro, declare it in your code before you declare `#include HLS/ac_int.h`.

**i++ Command Option Syntax**

```bash
-D DEBUG_AC_INT_ERROR
```

**Description**

Enables runtime tracking of ac_int data types during x86 emulation of your component (the `-march=x86-64` option, which is the default option, of the i++ command).

This tool uses additional resources to track the overflow and empty constructors, and emits a message for the first overflow that is detected and then exits the component with an error.

To learn more, review the tutorial: `<quartus_installdir>/hls/examples/tutorials/ac_datatypes/ac_int_overflow`
A. Advanced Math Source Code Libraries

The Intel HLS Compiler Pro Edition comes with templated source code libraries that help speed the development of your components by providing you with FPGA-optimized code for some commonly-used algorithms.

The Intel HLS Compiler provides the following libraries:

<table>
<thead>
<tr>
<th>Library</th>
<th>Description</th>
<th>Header file</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random number generator</td>
<td>Generate random integers or floating point numbers that follow a uniform distribution, or random floating point numbers that follow a Gaussian distribution</td>
<td>HLS/rand_lib.h</td>
</tr>
<tr>
<td>Matrix multiplication</td>
<td>Multiply two 2-D matrices.</td>
<td>HLS/matrix_mult.h</td>
</tr>
<tr>
<td>Cholesky decomposition</td>
<td>Factor a Hermitian (symmetric) positive-definite matrix into the product of a lower triangular matrix and its conjugate transpose $A = LL^T$.</td>
<td>HLS/cholesky_decompose.h</td>
</tr>
</tbody>
</table>

A.1. Random Number Generator Library

The random number generator source code library provided with the Intel HLS Compiler Pro Edition gives you FPGA-optimized random number generator template classes that you can add to your component without needing to write your own.

The Random Number Generator Library and Cryptography

The use of these pseudo-random number generator (PRNG) algorithms are not recommended for cryptographic purposes. The PRNGs included in this library are not cryptographically-secure pseudo-random number generators (CSPRNGs) and should not be used for cryptography. CSPRNG algorithms are designed so that no polynomial-time algorithm (PTA) can compute or predict the next bit in the pseudo-random sequence, nor is there a PTA that can predict past values of the CSPRNG; these algorithms do not achieve this purpose. Additionally, these algorithms have not been reviewed nor are they recommended for use as a PRNG component of a CSPRNG, even if the input values are from a non-deterministic entropy source with an appropriate entropy extractor.

Table 62. Properties of Values That Can Be Generated by the Intel HLS Compiler Random Number Generator Library

<table>
<thead>
<tr>
<th>Value distribution</th>
<th>Value type</th>
<th>Value range</th>
<th>Generation method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Uniform</td>
<td>Integer</td>
<td>$[-2^{31}, 2^{31}-1]$</td>
<td>Tausworthe Generator</td>
</tr>
<tr>
<td></td>
<td>Floating point</td>
<td>$(0, 1)$ (non-inclusive)</td>
<td>Tausworthe Generator</td>
</tr>
<tr>
<td>Gaussian</td>
<td>Floating point</td>
<td>$(0, 1)$</td>
<td>Central limit theorem (CLT) (Default)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Box-Muller</td>
</tr>
</tbody>
</table>
Header File

To include the random number generator library in your component, add the following line to your component:

```c
#include "HLS/rand_lib.h"
```

The header file is self-documented. You can review the header file to learn how to use the random number generator library in your component.

Random Number Object Declarations

Declare random number objects in your components as follows. In all cases, specifying `<seed_value>` is optional.

- Uniform distribution integer random number
  ```c
  static RNG_Uniform<int> <object_name>(<seed_value>)
  ```
- Uniform distribution floating point random number
  ```c
  static RNG_Uniform<float> <object_name>(<seed_value>)
  ```
- Gaussian distribution floating point random number (CLT method)
  ```c
  static RNG_Gaussian<float> <object_name>(<seed_value>)
  or
  static RNG_Gaussian<float, ihc::GAUSSIAN_CLT> <object_name>(<seed_value>)
  ```
- Gaussian distribution floating point random number (Box-Muller method)
  ```c
  static RNG_Gaussian<float, ihc::GAUSSIAN_BOX_MULLER> <object_name>(<seed_value>)
  ```

A.2. Matrix Multiplication Library

The matrix multiplication source code library provided with the Intel HLS Compiler Pro Edition gives you an FPGA-optimized templatized source code library to perform matrix multiplication of two matrices stored in a 2-D array.

When you use the matrix multiplication library, you can affect the number of DSP blocks and RAM blocks by controlling the dot product vector size and the number of matrix elements read at one time. Increasing the dot product vector size can achieve better latency, but at the cost of using more DSP blocks and other FPGA resources.

Header File

To include the matrix multiplication library in your component, add the following line to your component:

```c
#include "HLS/matrix_mult.h"
```

The header file is self-documented. You can review the header file to learn how to use the matrix multiplication library in your component.
Template Arguments

The matrix multiplication library multiplies two 2-D matrices, A and B. The resulting product is returned in a third matrix, C. The matrix multiplication library has the following template arguments:

- **T**
  - The data type of the matrix elements (For example, int, float, long, double).

- **t_rowsA**
  - The number of rows in matrix A.

- **t_colsA**
  - The number of columns in matrix A. This value also the number of rows in matrix B.

- **t_colsB**
  - The number of columns in matrix B.

- **DOT_VEC_SIZE**
  - The number of DSP blocks to use in a single computation. This value must be a factor of t_colsA.
  - You can achieve better component latency by increasing this value. However, you use more FPGA area to achieve this. Keeping this value low lowers your FPGA resource usage, but increases the latency.

- **BLOCK_SIZE**
  - The number of elements to read at one time from matrix A. The default value of BLOCK_SIZE is the value of DOT_VEC_SIZE. You can reduce this number if the bandwidth needed by matrix A is lower than the value of DOT_VEC_SIZE, but it must remain a factor of DOT_VEC_SIZE.

- **RUNNING_SUM_MULT_L**
  - This parameter can be adjusted to try and improve the fMAX of a component that uses this library. Review the header file for a detailed description of this argument and its effects.

A.3. Cholesky Decomposition Library

The Cholesky Decomposition library provided with the Intel HLS Compiler provides an FPGA-optimized templated library to factor a Hermitian (symmetric) positive-definite matrix into the product of a lower triangular matrix and its conjugate transpose \( A = LL^T \).

**Header File**

To include the Cholesky decomposition library in your component, add the following line to your component:

```
#include "HLS/cholesky_decompose.h"
```

The header file is self-documented. You can review the header file to learn how to use the Cholesky decomposition library. For high performance, use the -ffp-reassociate command flag when compiling your component.
**Variants**

You can call four different variants. All variants are in the `ihc::cholesky` namespace.

Cholesky decomposition iterates on the \( L \) matrix heavily. For the best component performance, the component need low latency access to the \( L \) matrix iterated on, and the \( L \) matrix must be vectorized. However, this requirement is not always easy to meet, so two of the variants allow you to use separate pieces of memory for the iterative decomposition and recording the final output.

The variants provided are as follows:

- `cholesky_decompose_real(A_input, L_iter, n);`
  
  Used for real-valued matrices.

- `cholesky_decompose_real(A_input, L_output, L_iter, n);`
  
  Used for real-valued matrices with separate memory for iterative decomposition and final output.

- `cholesky_decompose_complex(A_input, L_iter, n);`
  
  Used for complex-valued matrices.

- `cholesky_decompose_complex(A_input, L_output, L_iter, n);`
  
  Used for complex-valued matrices with separate memory for iterative decomposition and final output.

**Arguments**

- **A_input**
  
  A pointer to an array of size `MATRIX_SIZE * MATRIX_SIZE`, for providing input \( A \) matrix.

  Its data width and alignment should match that of a single element in the matrix.

- **L_iter**
  
  A pointer to an array of size `MATRIX_SIZE * MATRIX_SIZE`, for holding the \( L \) matrix that the program can iterate on.

  In the 3-argument variants, this is the same place for recording the result. Its data width and alignment should match that of a vector of elements in the matrix, with specified vectorization width.

  If a pointer to component memory is given, the compiler should be able to optimize that automatically.

- **L_output**
  
  A pointer to an array of size `MATRIX_SIZE * MATRIX_SIZE`, for recording the final \( L \) matrix result in the 4-argument variants.

  Its data width and alignment should match that of a single element in the matrix.

- **n**
  
  The actual matrix size.

  It should not be greater than `MATRIX_SIZE`
Template Arguments

**FP_T**  
The floating-point data type. Can be *float* or *double*.

**VEC_SIZE_PWR**  
The logarithm (base 2) of the vectorization width. Vectorization with a width of $2^{VEC_SIZE_PWR}$ is used for the access of L_iter matrix and the dot product computation.

This header uses the Intel HLS Compiler ivdep pragma with its `safelen()` clause, Dummy iterations matching the `safelen()` values are inserted to manage memory access dependency more efficiently.

The ivdep pragma with its `safelen()` clause are used for the following template arguments in the Cholesky decomposition algorithm:

- **INNER_SAFELEN** value applies to the partial_dot array,
- **OUTER_SAFELEN_OVERWRITE** value applies to the L_iter matrix

Although a reasonable estimate of the `safelen()` value is provided, their value might require some tuning for different devices, clock targets, precisions, and memory arrangements.

Use an iterative approach for finding their optimal values. First, try large conservative estimates until the compilation result does not demonstrate any II issues and gives you a a satisfactory $f_{\text{MAX}}$. Then, use the Function Viewer (part of the High-Level Design Reports) to examine the schedule of the load and store nodes for the memory the `safelen()` value applies to. The difference of start cycle of the load and store should approximately match the optimal value to use for that `safelen()` value.

**INNER_SAFELEN**  
The INNER_SAFELEN value applies to the partial_dot array. Its value roughly matches the latency of index calculation and a floating point addition in the algorithm. This `safelen()` value does not grow with matrix size. A default value of 16 is given, which should be ideal for single precision. For double precision, increase it accordingly.

**OUTER_SAFELEN_OVERWRITE**  
The OUTER_SAFELEN_OVERWRITE value applies to the L_iter matrix, and is related to intercolumn dependencies. An estimate of this `safelen()` value is provided and along with an estimation of how this value grows along with matrix size. However, the optimal value for this template argument depends on clock target values, target device families, and the precision used.

Use this parameter to overwrite the set `safelen()` value.

If **OUTER_SAFELEN_OVERWRITE=-1**, the `safelen()` value is tuned for single-precision float data types on an Intel Arria 10 device with a default clock.
target. The tuning also assumes that the $L_{\text{iter}}$ matrix is in component memory with proper vectorization provided.
B. Supported Math Functions

The Intel HLS Compiler has built-in support for generating efficient IP out of standard math functions present in the math.h C header file. The compiler also has support for some math functions that are not supported by the math.h header file, and these functions are provided in extendedmath.h C header file.

To use the Intel implementation of math.h for Intel FPGAs, include HLS/math.h in your function by adding the following line:

```cpp
#include "HLS/math.h"
```

To use the nonstandard math functions that are optimized for Intel FPGAs, include HLS/extendedmath.h in your function by adding the following line:

```cpp
#include "HLS/extendedmath.h"
```

The extendedmath.h header is compatible only with Intel HLS Compiler. It is not compatible with GCC or Microsoft Visual Studio.

If your component uses arbitrary precision fixed-point datatypes provided in the ac_fixed.h header, you use some of the datatypes with some math functions by including the following line:

```cpp
#include "HLS/ac_fixed_math.h"
```

To see examples of how to use the math functions provided by these header files, review the following tutorial: `<quartus_installdir>/hls/examples/tutorials/best_practices/single_vs_double_precision_math`.

If your component uses the hls_float arbitrary precision floating point data type, add the following line to add support for math functions:

```cpp
#include "HLS/hls_float_math.h"
```

B.1. Math Functions Provided by the math.h Header File

The Intel HLS Compiler Pro Edition supports a subset of functions that are present in your native compiler through the HLS/math.h header file.

For each math.h function listed below, "●" indicates that the HLS compiler supports the function; "X" indicates that the function is not supported.

The math functions supported on Linux operating systems might differ from the math functions supported on Windows operating systems. Review the comments in the HLS/math.h header file to see which math functions are supported on the different operating systems.
### Trigonometric Functions

<table>
<thead>
<tr>
<th>Trigonometric Function</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>cos</td>
<td>cosf</td>
</tr>
<tr>
<td>sin</td>
<td>sinf</td>
</tr>
<tr>
<td>tan</td>
<td>tanf</td>
</tr>
<tr>
<td>acos</td>
<td>acosf</td>
</tr>
<tr>
<td>asin</td>
<td>asinf</td>
</tr>
<tr>
<td>atan</td>
<td>atanf</td>
</tr>
<tr>
<td>atan2</td>
<td>atan2f</td>
</tr>
</tbody>
</table>

### Hyperbolic Functions

<table>
<thead>
<tr>
<th>Hyperbolic Function</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>cosh</td>
<td>coshf</td>
</tr>
<tr>
<td>sinh</td>
<td>sinhf</td>
</tr>
<tr>
<td>tanh</td>
<td>tanhf</td>
</tr>
<tr>
<td>acosh</td>
<td>acoshf</td>
</tr>
<tr>
<td>asinh</td>
<td>asinhf</td>
</tr>
<tr>
<td>atanh</td>
<td>atanhf</td>
</tr>
</tbody>
</table>

### Exponential and Logarithmic Functions

<table>
<thead>
<tr>
<th>Exponential or Logarithmic Function</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>exp</td>
<td>expf</td>
</tr>
<tr>
<td>frexp</td>
<td>frexp</td>
</tr>
<tr>
<td>ldexp</td>
<td>ldexpf</td>
</tr>
<tr>
<td>log</td>
<td>logf</td>
</tr>
<tr>
<td>log10</td>
<td>log10f</td>
</tr>
<tr>
<td>modf</td>
<td>modff</td>
</tr>
<tr>
<td>exp2</td>
<td>exp2f</td>
</tr>
<tr>
<td>exp10 (Linux only)</td>
<td>exp10f (Linux only) (*)</td>
</tr>
<tr>
<td>expm1</td>
<td>expm1f</td>
</tr>
<tr>
<td>ilogb</td>
<td>ilogbf</td>
</tr>
<tr>
<td>log1pf</td>
<td>log1pf</td>
</tr>
<tr>
<td>log2</td>
<td>log2f</td>
</tr>
</tbody>
</table>

(*) For Windows, support for this function is in the `extendedmath.h` header file.
### Exponential or Logarithmic Function

<table>
<thead>
<tr>
<th></th>
<th>Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>logb</td>
<td></td>
</tr>
<tr>
<td>scalbn</td>
<td></td>
</tr>
<tr>
<td>scalbin</td>
<td></td>
</tr>
</tbody>
</table>

### Power Functions

<table>
<thead>
<tr>
<th>Power Function</th>
<th>Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>pow</td>
<td></td>
</tr>
<tr>
<td>sqrt</td>
<td></td>
</tr>
<tr>
<td>cbrt</td>
<td></td>
</tr>
<tr>
<td>hypot</td>
<td></td>
</tr>
</tbody>
</table>

### Error and Gamma Functions

<table>
<thead>
<tr>
<th>Error or Gamma Function</th>
<th>Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>erf</td>
<td></td>
</tr>
<tr>
<td>erfc</td>
<td></td>
</tr>
<tr>
<td>tgamma</td>
<td></td>
</tr>
<tr>
<td>lgamma</td>
<td></td>
</tr>
<tr>
<td>lgamma_r (Linux only)</td>
<td></td>
</tr>
</tbody>
</table>

### Rounding and Remainder Functions

<table>
<thead>
<tr>
<th>Rounding or Remainder Function</th>
<th>Supported</th>
</tr>
</thead>
<tbody>
<tr>
<td>ceil</td>
<td></td>
</tr>
<tr>
<td>floor</td>
<td></td>
</tr>
<tr>
<td>fmod</td>
<td></td>
</tr>
<tr>
<td>trunc</td>
<td></td>
</tr>
<tr>
<td>round</td>
<td></td>
</tr>
<tr>
<td>lround</td>
<td></td>
</tr>
<tr>
<td>lroundf</td>
<td></td>
</tr>
<tr>
<td>rint</td>
<td></td>
</tr>
<tr>
<td>lrint</td>
<td></td>
</tr>
<tr>
<td>ilrint</td>
<td></td>
</tr>
<tr>
<td>nearbyint</td>
<td></td>
</tr>
<tr>
<td>remainder</td>
<td></td>
</tr>
<tr>
<td>remainderf</td>
<td></td>
</tr>
<tr>
<td>remquo</td>
<td></td>
</tr>
</tbody>
</table>
### Table 69. Floating-Point Manipulation Functions

<table>
<thead>
<tr>
<th>Floating-Point Manipulation Function</th>
<th>Supported ?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td>Single-precision floating point function</td>
</tr>
<tr>
<td>copysign</td>
<td>copysignf</td>
</tr>
<tr>
<td>nan</td>
<td>nanf</td>
</tr>
<tr>
<td>nextafter</td>
<td>nextafterf</td>
</tr>
<tr>
<td>nexttoward</td>
<td>nexttowardf</td>
</tr>
</tbody>
</table>

### Table 70. Minimum, Maximum, and Difference Functions

<table>
<thead>
<tr>
<th>Minimum, Maximum, or Difference Function</th>
<th>Supported ?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td>Single-precision floating point function</td>
</tr>
<tr>
<td>fdim</td>
<td>fdim</td>
</tr>
<tr>
<td>fmax</td>
<td>fmax</td>
</tr>
<tr>
<td>fmin</td>
<td>fmin</td>
</tr>
</tbody>
</table>

### Table 71. Other Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Supported ?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td>Single-precision floating point function</td>
</tr>
<tr>
<td>fabs</td>
<td>fabsf</td>
</tr>
<tr>
<td>fma</td>
<td>fmaf</td>
</tr>
</tbody>
</table>

### Table 72. Classification Macros

<table>
<thead>
<tr>
<th>Classification Macro</th>
<th>Supported ?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td>Single-precision floating point function</td>
</tr>
<tr>
<td>fpclassify (Linux only)</td>
<td>fpclassifyf (Linux only)</td>
</tr>
<tr>
<td>isfinite</td>
<td>isnfinite</td>
</tr>
<tr>
<td>isninf</td>
<td>isninff</td>
</tr>
<tr>
<td>isnan</td>
<td>isnanf</td>
</tr>
<tr>
<td>isnormal (Linux only)</td>
<td>isnormalf (Linux only)</td>
</tr>
<tr>
<td>signbit (Linux only)</td>
<td>signbitf (Linux only)</td>
</tr>
</tbody>
</table>

### Table 73. Comparison Macros

<table>
<thead>
<tr>
<th>Comparison Macro</th>
<th>Supported ?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td>Single-precision floating point function</td>
</tr>
<tr>
<td>isgreater</td>
<td>isgreaterf</td>
</tr>
<tr>
<td>isgreaterequal</td>
<td>isgreaterequalf</td>
</tr>
<tr>
<td>isless</td>
<td>islessf</td>
</tr>
</tbody>
</table>

*continued...*
B.2. Math Functions Provided by the `extendedmath.h` Header File

The Intel HLS Compiler Pro Edition supports an additional subset of math functions through the `HLS/extendedmath.h` header file.

For each `extendedmath.h` function listed below, "●" indicates that the Intel HLS Compiler Pro Edition supports the function; "X" indicates that the function is not supported.

The math functions supported on Linux operating systems might differ from the math functions supported on Windows operating systems. Review the comments in the `HLS/extendedmath.h` header file to see which math functions are supported on the different operating systems.

Table 74. Extended Math Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Supported?</th>
<th>Double-precision floating point function</th>
<th>Single-precision floating point function</th>
</tr>
</thead>
<tbody>
<tr>
<td>sincos</td>
<td>●</td>
<td>sincosf</td>
<td></td>
</tr>
<tr>
<td>acospi</td>
<td>●</td>
<td>acospif</td>
<td></td>
</tr>
<tr>
<td>asinpi</td>
<td>●</td>
<td>asinpif</td>
<td></td>
</tr>
<tr>
<td>atanpi</td>
<td>●</td>
<td>atanpif</td>
<td></td>
</tr>
<tr>
<td>cospi</td>
<td>●</td>
<td>cospif</td>
<td></td>
</tr>
<tr>
<td>sinpi</td>
<td>●</td>
<td>sinpif</td>
<td></td>
</tr>
<tr>
<td>tanpi</td>
<td>●</td>
<td>tanpif</td>
<td></td>
</tr>
<tr>
<td>pown</td>
<td>●</td>
<td>pownf</td>
<td></td>
</tr>
<tr>
<td>powr</td>
<td>●</td>
<td>powrf</td>
<td></td>
</tr>
<tr>
<td>rsqrt</td>
<td>●</td>
<td>rsqrdf</td>
<td></td>
</tr>
</tbody>
</table>

Table 75. Exponential and Logarithmic Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Supported?</th>
<th>Double-precision floating point function</th>
<th>Single-precision floating point function</th>
</tr>
</thead>
<tbody>
<tr>
<td>exp10 (Windows only)</td>
<td>●</td>
<td>exp10f (Windows only) (*)</td>
<td></td>
</tr>
</tbody>
</table>

(*) For Linux, support for this function is in the `math.h` header file
### Table 76. Error and Gamma Functions

<table>
<thead>
<tr>
<th>Error or Gamma Function</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td></td>
</tr>
<tr>
<td>Single-precision floating point function</td>
<td></td>
</tr>
<tr>
<td>lgamma_r (Windows only) (*)</td>
<td>●</td>
</tr>
<tr>
<td>lgamma_rf (Windows only) (*)</td>
<td>●</td>
</tr>
</tbody>
</table>

### Table 77. Minimum, Maximum, and Difference Functions

<table>
<thead>
<tr>
<th>Minimum, Maximum, or Difference Function</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td></td>
</tr>
<tr>
<td>Single-precision floating point function</td>
<td></td>
</tr>
<tr>
<td>maxmag</td>
<td>●</td>
</tr>
<tr>
<td>minmag</td>
<td>●</td>
</tr>
</tbody>
</table>

### Table 78. Other Functions

<table>
<thead>
<tr>
<th>Function</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td></td>
</tr>
<tr>
<td>Single-precision floating point function</td>
<td></td>
</tr>
<tr>
<td>fract</td>
<td>●</td>
</tr>
<tr>
<td>mad</td>
<td>●</td>
</tr>
<tr>
<td>oclnan</td>
<td>●</td>
</tr>
<tr>
<td>rootn</td>
<td>●</td>
</tr>
</tbody>
</table>

### Table 79. Classification Macros

<table>
<thead>
<tr>
<th>Classification Macro</th>
<th>Supported?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Double-precision floating point function</td>
<td></td>
</tr>
<tr>
<td>Single-precision floating point function</td>
<td></td>
</tr>
<tr>
<td>isordered</td>
<td>●</td>
</tr>
</tbody>
</table>

In addition, the HLS/extendedmath.h header file supports the following versions of the popcount function:

### Table 80. Popcount function

<table>
<thead>
<tr>
<th>Data type</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unsigned char</td>
<td>popcountc</td>
</tr>
<tr>
<td>Unsigned short</td>
<td>popcounts</td>
</tr>
<tr>
<td>Unsigned int</td>
<td>popcount</td>
</tr>
<tr>
<td>Unsigned long</td>
<td>popcountl</td>
</tr>
<tr>
<td>Unsigned long long</td>
<td>popcountll</td>
</tr>
</tbody>
</table>

To see an example of how to use the math functions provided by the extendedmath.h header file and how to override a math function in the header file so that you can compile your design with GCC or Microsoft Visual Studio, review the following example design: <quartus_installdir>/hls/examples/QRD.
B.3. Math Functions Provided by the `ac_fixed_math.h` Header File

Adding the `ac_fixed_math.h` header file adds support for the following arbitrary precision fixed-point (`ac_fixed`) datatype functions:

- `sqrt_fixed`
- `reciprocal_fixed`
- `reciprocal_sqrt_fixed`
- `sin_fixed`
- `cos_fixed`
- `sincos_fixed`
- `sinpi_fixed`
- `cospi_fixed`
- `sincospi_fixed`
- `log_fixed`
- `exp_fixed`

For details about inputs type restrictions, input value limits, and output type propagation rules, review the comments in the `ac_fixed_math.h` header file.

*Important:* Due to the differences in the internal math implementations, the results from operations with `ac_fixed` data types might be different between simulation and emulation. The maximum difference will be within a few ULPs.

B.4. Math Functions Provided by the `hls_float.h` Header File

Adding the `hls_float.h` header file adds support for the following arbitrary precision floating point (`hls_float.h`) data type functions:

- Arithmetic operators: `+`, `-`, `*`, `/`
- Arithmetic assignment operators: `=`, `+=`, `-=`, `*=`
- Comparison operators: `>`, `<`, `>=`, `<=`
- Unary operators: `+()`, `-()`, `abs()` 
- Explicit functions: `add(a, b)`, `sub(a, b)`, `mul(a, b)`, `div(a, b)`

B.5. Math Functions Provided by the `hls_float_math.h` Header File

Adding the `hls_float_math.h` header file adds support for the following arbitrary precision floating point (`hls_float_math.h`) data type functions:

- square root: `ihc_sqrt`
- cube root: `ihc_cbrt`
- reciprocal (inverse): `ihc_recip`
- reciprocal (inverse) square root: `ihc_rsqrt`
- hypotenuse: `ihc_hypot`
B. Supported Math Functions

- $e^x$: ihc_exp
- $e^{x-1}$: ihc_expm1
- $2^x$: ihc_exp2
- $10^x$: ihc_exp10
- $\ln(x)$: ihc_log
- $\log_2(x)$: ihc_log2
- $\log_{10}(x)$: ihc_log10
- $\ln(1+x)$: ihc_log1p
- $x^y$: ihc_pow
  Both $x$ and $y$ are hls_float data types
- $x^n$: ihc_pown
  $x$ is hls_float data type, $n$ is ac_int data type
- $x^y$: ihc_powr
  Both $x$ and $y$ are hls_float data types. With the restriction that $x>0 \rightarrow x^y$, $x>0$, undefined behavior if $x < 0$
- $\sin$: ihc_sin
- $\sin\pi$: ihc_sinpi
- $\cos$: ihc_cos
- $\cos\pi$: ihc_cospi
- $\sin\cos$: ihc_sincos
- $\sin\cos\pi$: ihc_sincospi
- $\arcsin$: ihc_asin
- $\arcsin\pi$: ihc_asinpi
- $\arccos$: ihc_acos
- $\arccos\pi$: ihc_acospi
- $\arctan$: ihc_atan
- $\arctan\pi$: ihc_atanpi
- $\arctan(x/y)$: ihc_atan2
  Both $x$ and $y$ are hls_float data types

For details about inputs type restrictions, input value limits, and output type propagation rules, review the comments in the hls_float_math.h header file.

B.6. Default Rounding Schemes and Subnormal Number Support

For some floating-point operations defined in the math.h and extendedmath.h header files, you can control whether the operations should support subnormal numbers with the --daz i++ command option. Disabling subnormal number support can save FPGA area by rounding subnormal numbers to zero.
In addition, you can control the rounding scheme used for single- and double-
precision adders, multipliers, and dividers with the --rounding i++ command
option.

For single-precision operation, subnormal support does not apply.

The following tables show the default subnormal support and rounding scheme for
single- and double-precision floating point operators.

Table 81. Default Subnormal Support and Rounding Scheme for Single-Precision
Floating Point Operators

For descriptions of the rounding schemes, see the --rounding i++ command option description in Intel HLS

<table>
<thead>
<tr>
<th>Single-Precision Floating-Point Operator</th>
<th>Default Subnormal Support</th>
<th>Default Rounding Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>N/A</td>
<td>IEEE-754 RNE</td>
</tr>
<tr>
<td>-</td>
<td>N/A</td>
<td>IEEE-754 RNE</td>
</tr>
<tr>
<td>*</td>
<td>N/A</td>
<td>IEEE-754 RNE</td>
</tr>
<tr>
<td>/</td>
<td>N/A</td>
<td>Faithful</td>
</tr>
</tbody>
</table>

Table 82. Default Subnormal Support and Rounding Scheme for Double-Precision
Floating Point Operators

For descriptions of the rounding schemes, see the --rounding i++ command option description in Intel HLS

<table>
<thead>
<tr>
<th>Double-Precision Floating-Point Operator</th>
<th>Default Subnormal Support</th>
<th>Default Rounding Scheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>On</td>
<td>IEEE-754 RNE</td>
</tr>
<tr>
<td>-</td>
<td>On</td>
<td>IEEE-754 RNE</td>
</tr>
<tr>
<td>*</td>
<td>On</td>
<td>IEEE-754 RNE</td>
</tr>
<tr>
<td>/</td>
<td>On</td>
<td>Faithful</td>
</tr>
</tbody>
</table>

For the following double-precision floating-point functions, subnormal support is on by
default:

- sqrt
- floor
- ceil
- cbrt
- rsqrt
- hypot
- modf
- exp
- expm1
- exp2
- exp10
- log
- log1p
- log2
- log10
- pow
- powr
- sin
- cos
- tan
- sinpi
- cospi

The rounding schemes do not apply to these functions.
Cyclone V Restrictions

Using the Intel HLS Compiler Pro Edition to compile designs that target the Cyclone V device family is subject to a number of restrictions.

- Cyclone V devices do not have hardened floating-point DSP blocks. Designs that target Cyclone V devices use soft-logic for DSP functions such as multiplication, addition, or square root.

The Intel HLS Compiler Pro Edition cannot infer hardened floating-point dot products for Cyclone V devices, but the compiler can rearrange operations to improve latency if you specify the `-ffp-reassociate` and/or `-ffp-contract=fast` options of the `i++` command.

Archives

<table>
<thead>
<tr>
<th>Intel HLS Compiler Version</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>19.3</td>
<td>Intel HLS Compiler Reference Manual</td>
</tr>
<tr>
<td>19.2</td>
<td>Intel HLS Compiler Reference Manual</td>
</tr>
<tr>
<td>19.1</td>
<td>Intel HLS Compiler Reference Manual</td>
</tr>
<tr>
<td>18.1.1</td>
<td>Intel HLS Compiler Reference Manual</td>
</tr>
<tr>
<td>18.1</td>
<td>Intel HLS Compiler Reference Manual</td>
</tr>
<tr>
<td>18.0</td>
<td>Intel HLS Compiler Reference Manual</td>
</tr>
<tr>
<td>17.1.1</td>
<td>Intel HLS Compiler Reference Manual</td>
</tr>
<tr>
<td>17.1</td>
<td>Intel HLS Compiler Reference Manual</td>
</tr>
</tbody>
</table>

|------------------|----------------------------------------|---------|
| 2021.05.21       | 21.2                                   | • Added limitation for casting `hls_float` datatypes to unsigned integer datatypes to Arbitrary Precision Math Support on page 79.  
• Revised references to Avalon Interfaces to align with new Avalon Interconnect terminology. Avalon master interfaces are now Avalon host interfaces, and Avalon slave interfaces are now Avalon agent interfaces.  
• Renamed the interfaces to align with the new Avalon Interconnect terminology. Interfaces using the old terminology are deprecated. Interfaces using the new terminology have been introduced as replacements as follows:

<table>
<thead>
<tr>
<th>Deprecated Interface</th>
<th>Replacement Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>hls_avalon_slave_component</code></td>
<td><code>hls_avalon_agent_component</code></td>
</tr>
<tr>
<td><code>hls_avalon_slave_register_argument</code></td>
<td><code>hls_avalon_agent_register_argument</code></td>
</tr>
<tr>
<td><code>hls_avalon_slave_memory_argument</code></td>
<td><code>hls_avalon_agent_memory_argument</code></td>
</tr>
<tr>
<td><code>mm_master</code></td>
<td><code>mm_host</code></td>
</tr>
</tbody>
</table>

If you use the deprecated interfaces, you might get warning messages when you compile your component.

| 2021.03.29       | 21.1                                   | • Added the following topics about hardware implementation control for math functions:  
— Hardware Implementation Control for Math Functions on page 139  
— The `--dsp-mode` Command Option on page 141  
— The `ihc::math_dsp_control` Function on page 141  
— Added the `--dsp-mode` command option to the following topics:  
  • Intel HLS Compiler Pro Edition Command Options on page 6  
  • Intel HLS Compiler Pro Edition i++ Command-Line Arguments on page 144  
— Added the `math_dsp_control.h` header file to the following topics:  
  • C and C++ Libraries on page 16  
  • Intel HLS Compiler Pro Edition Header Files on page 147  
• Updated Task Functions on page 101 to include the following new sections:  
  • Task Invocation Interface on page 102  
  • Task Capacity on page 108  
• Added Compiling Your Testbench and Component Separately on page 14.  
  • Also, moved some content from Compiler Interoperability on page 11 to Compiling Your Testbench and Component Separately on page 14.  
• Deprecated the `--ffp-reassoc` command option in the following topics. Use the `--ffp-reassociate` command option instead.  
  • Intel HLS Compiler Pro Edition Command Options on page 6  
  • Intel HLS Compiler Pro Edition i++ Command-Line Arguments on page 144

continued...
|------------------|-----------------------------------|---------|
|                  |                                   | • Revised the following topics to refer to the `-ffp-reassociate` command option instead of the `-ffp-reassoc` command option:  
  — Cholesky Decomposition Library on page 200  
  — Cyclone V Restrictions on page 213  
• Deprecated the `fp reassoc` scoped pragma in [Intel HLS Compiler Pro Edition Scope Pragmas](#) on page 166. Use the `fp reassociate` scoped pragma instead.  
• Replaced occurrences of "Graph Viewer" with "System Viewer". |
| 2020.12.14       | 20.4                              | • Added documentation for Cyclone V device support:  
  — Added `CycloneV` to the values accepted by the `-march` compiler option in [Intel HLS Compiler Pro Edition Command Options](#) on page 6.  
  — Added list of restrictions to [Cyclone V Restrictions](#) on page 213.  
Cyclone V device support requires Intel Quartus Prime Standard Edition. For details, see [Intel HLS Compiler Pro Edition Getting Started Guide](#).  
• Moved Pipes on page 28 from being a separate chapter to being part of the Component Interfaces on page 21 chapter. |
| 2020.09.28       | 20.3                              | • Corrected the syntax of the `-ghdl1++` command option in [Intel HLS Compiler Pro Edition Command Options](#) on page 6 and [Intel HLS Compiler Pro Edition i++ Command-Line Arguments](#) on page 144.  
The correct syntax is `-ghdl[=<depth>]`.  
• Removed the `--fpc i++` command option. Use the `-ffp-contract=fast` command option instead.  
• Removed the `--fp-relaxed i++` command option. Use `-ffp-reassoc` command option instead.  
• Updated [unroll Loop Pragma](#) on page 165 with warning that unrolling nested loops with large bounds might cause very long compile times.  
• Updated limitations on dividers for AC data types in [Arbitrary Precision Math Support](#) on page 79.  
• Added additional information about task function interfaces to [Task Functions](#) on page 101.  
• Added an example to illustrate how the compiler transforms your code when unrolling loops with `#pragma unroll` in [Loop Unrolling (unroll Pragma)](#) on page 65.  
• Removed the restriction that loops considered for fusing must have the same trip count. The compiler supports fusing loops with different trip counts.  
• Updated [Loop Fusion](#) on page 71 and [Loop Fusion Control (loop_fuse Pragma)](#) on page 73 with information about fusing loops with different trip counts.  
• Updated [Default Rounding Schemes and Subnormal Number Support](#) on page 211 with information for single-precision floating point operations. |
| 2020.06.22       | 20.2                              | • Updated content for Intel Agilex device support.  
• Updated [Component Invocation Interface](#) on page 21 to show block diagram of default component interface.  
• Added [Interface Definition Example: Component Invocation Interface Control Attributes](#) on page 23 to show additional example block diagrams of component interfaces  
• Updated [Loop Interleaving Control (max_interleaving Pragma)](#) on page 68 with an example showing the difference between regular pipelined loop execution and interleaved loop execution.  
• Updated [Integer Promotion and ac_int Data Types](#) on page 82 to include a description of how using `ac_int` data types affects the data width of operations.  
• Added [Additional Data Types Provided By hls_float.h](#) on page 97 to describe the `bfloat16` and `bfloat19` aliases for `hls_float` data types, and the FPGA hardware implications of implementing these data types.  
• Updated the description of `ihc::launch_always_run` on page 99 to include a recommendation to use `ihc_hls_set_component_wait_cycle` with it to keep the component and always-run tasks correctly coordinated.  
• Added [Effects of Specifying Target II and Target fMAX](#) on page 99 to describe the interactions of specifying a target II, a target fMAX, or both. |
|------------------|--------------------------------------|---------|
| 2020.05.13       | 20.1                                 | • Revised description of the ihc::stream template object in ihc::stream Class on page 179 to specify that the template object can take only trivially copyable C++ data types. Not all valid C++ data types are trivially copyable.  
• Updated Slave Memories with a diagram illustrating slave memory interfaces.  
• Updated the following topics to describe the wait time implemented by the ihc_hls_set_component_wait_cycle simulation API function:  
  — System of Tasks Simulation on page 110  
  — Intel HLS Compiler Pro Edition Simulation API (Testbench Only) on page 151  
• Updated Intel HLS Compiler Pro Edition Systems of Tasks API on page 176 with additional information about when to specify a value for the optional capacity parameter.  
• Moved the example previously in Loop Initiation Interval (ii Pragma) on page 62 into Raise Loop II to Increase $f_{MAX}$ in the Intel HLS Compiler Best Practice Guide.  
• Fixed a link in Component Target Frequency on page 98. The --clock i++ command option affects compilation and not linking.  
• Corrected a typo in Compiler Interoperability on page 11. The correct option for generating debug symbols with the g++ command is -g. |
• Added Default Rounding Schemes and Subnormal Number Support on page 211.  
• Added the hls_private_copies memory attribute to Intel HLS Compiler Pro Edition Component Memory Attributes on page 153.  
• Deprecated the hls_max_concurrency memory attribute in Intel HLS Compiler Pro Edition Component Memory Attributes on page 153. Use the hls_private_copies memory attribute instead.  
• Added hls_use_stall_enabled_clusters component attribute to Intel HLS Compiler Pro Edition Component Attributes on page 168.  
• Updated Static Variables on page 58 to indicate that static variables cannot be shared by different distinct task or component function. They can be shared by different invocation of the same function.  
• Added or revised the following sections to add information about loop fusion:  
  — Loop Fusion on page 71  
  — Loop Fusion Control (loop_fuse Pragma) on page 73  
  — Loop Fusion Exemption (nofusion pragma) on page 75  
  — Intel HLS Compiler Pro Edition Loop Pragmas  
• Added information about stream support for RTL objects in object libraries to the following sections:  
  — RTL Module Interface Signals on page 120  
  — Mapping HLS Data Types to RTL Signals on page 131  
  — XML Elements for INTERFACE on page 129  
• Added a restriction for creating creating objects from HLS code targeted for use with the Intel FPGA SDK for OpenCL in Creating Objects From HLS Code on page 114.  
• Updated description of ALLOW_MERGING element in XML Elements for ATTRIBUTES on page 126 and added Interaction Between ALLOW_MERGING and HAS_SIDE_EFFECTS Elements on page 128.  
• Updated description of the -L i++ command option to indicate that you can have a space (or not) between the option and the specified search path.  
• Updated the following topics to document the discrepancy between simulation and emulation results when using ac_fixed_math.h provided math functions:  
  — Declaring ac_fixed Data Types on page 86  
  — Math Functions Provided by the ac_fixed_math.h Header File on page 210  
• Added the hls_force_pow2_depth memory attribute to Intel HLS Compiler Pro Edition Component Memory Attributes on page 153.  
• Updated the list of restrictions on HLS task functions in Task Functions on page 101. |

...continued...
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>• Removed the <code>hls_numports_readonly_writeonly</code> memory attribute. It is no longer supported. Use <code>hls_max_replicates</code> instead.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Added information about using the <code>volatile</code> keyword with slave memories to <code>Slave Memories</code>.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Added information about the <code>hls_readwrite_mode</code> macro to the following topics:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— Agent Memories on page 46</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— Intel HLS Compiler Pro Edition Component Macros on page 173</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Updated the syntax of the <code>ihc:launch</code> and <code>ihc:collect</code> functions in the following topics:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— Intel HLS Compiler Pro Edition Systems of Tasks API on page 176</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— Templated Functions on page 18</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Added the optional <code>capacity</code> parameter to the <code>ihc:launch</code> and <code>ihc:collect</code> functions in Intel HLS Compiler Pro Edition Systems of Tasks API on page 176.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Added a restriction on loops in Supported C and C++ Subset for Component Synthesis on page 16.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Revised the requirements when using arbitrary precision datatypes in libraries in Supported OpenCL Language Constructs on page 116.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Added Cholesky Decomposition Library on page 200 to Advanced Math Source Code Libraries on page 198.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Added information about using pipes in your component and task functions to the following sections:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— Pipes on page 28</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— The pipe Class and Its Use on page 30</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— C and C++ Libraries on page 16</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— Intel HLS Compiler Pro Edition Header Files on page 147</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Moved information from Component Invocation Interface Control Attributes section into Component Invocation Interface on page 21.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• In XML Elements for INTERFACE on page 129, removed restriction on <code>width</code> attribute of the <code>INPUT</code> element. Input ports are not restricted to widths that are multiples of 8 bits. They can have any arbitrary width</td>
</tr>
</tbody>
</table>

2020.02.10 19.4 • Corrected the syntax of the `ihc::launch_always_run` system of tasks API function in Intel HLS Compiler Pro Edition Systems of Tasks API on page 176.

2020.01.27 19.4 • Corrected the spelling of the `-ffp-contract=fast` and `-ffp-reassoc` command options in the following sections: |
|                  |                                       |   — Intel HLS Compiler Pro Edition Command Options on page 6 |
|                  |                                       |   — Intel HLS Compiler Pro Edition i++ Command-Line Arguments on page 144 |
|                  |                                       |   — Intel HLS Compiler Pro Edition Scope Pragmas on page 166 |

|                  |                                       | • Added `ihc::launch_always_run` to Intel HLS Compiler Pro Edition Systems of Tasks API on page 176. |
|                  |                                       | • Added descriptions of `--daz` and `--rounding` command options to Intel HLS Compiler Pro Edition Command Options on page 6 and Intel HLS Compiler Pro Edition i++ Command-Line Arguments on page 144 |
|                  |                                       | • Updated Operators and Return Types Supported by the `hls_float Data Type` on page 92 with addition details about type conversion to and from the `hls_float` data type. |
|                  |                                       | • Added Creating a Static-Object Library on page 113. |
|                  |                                       | • Updated Object Manifest File Syntax on page 124 with information about `resetn` signal handling. |


<table>
<thead>
<tr>
<th>Document Version</th>
<th>Intel Quartus Prime Version</th>
<th>Changes</th>
</tr>
</thead>
<tbody>
<tr>
<td>2019.09.30</td>
<td>19.3</td>
<td>• <strong>PRO</strong> Expanded and reorganized information about HLS libraries into a new chapter that starts with <em>Static-Object Libraries</em> on page 112.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Added information about arbitrary precision floating point number support to <em>Declaring hls_float Data Types</em> on page 90.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Added <em>Intel HLS Compiler Pro Edition Scope Pragmas</em> on page 166.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> For variable-latency Avalon Memory-Mapped (MM) Master interfaces, added information about load-store unit control to <em>Intel HLS Compiler Pro Edition Load-Store Unit Control</em> on page 194.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Added the <strong>--ffp-reassoc</strong> and <strong>--ffp-contract=fast</strong> options to <em>Intel HLS Compiler Pro Edition i++ Command-Line Arguments</em> on page 144.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Revised <em>Component Memories (Memory Attributes)</em> on page 52 (formerly <em>Local Variables in Components</em>).</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Added information about the <em>hls_max_replicates</em> memory attribute to the following sections:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— <em>Component Memories (Memory Attributes)</em> on page 52</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— <em>Intel HLS Compiler Pro Edition Component Memory Attributes</em> on page 153</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Deprecated the <em>hls_numports readonlyonly writeonly</em> memory attribute throughout this document. Use <em>hls_max_replicates</em> instead.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Added information about the <em>max_interleaving</em> loop pragma to the following sections:</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— <em>Loop Interleaving Control (max_interleavingPragma)</em> on page 68</td>
</tr>
<tr>
<td></td>
<td></td>
<td>— <em>Intel HLS Compiler Pro Edition Loop Pragmas</em> on page 161</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Added information about the <em>hls_fpga_reg</em> function to <em>Advanced Hardware Synthesis Controls</em> on page 139.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• <strong>PRO</strong> Removed the restriction that task functions cannot have pointer or reference arguments from <em>Task Functions</em> on page 101.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• In <em>Intel HLS Compiler Pro Edition Component Memory Attributes</em> on page 153, revised the description of the default value of the <em>hls_bankbits</em> memory attribute.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• In <em>Intel HLS Compiler Pro Edition Component Memory Attributes</em> on page 153, removed references to the <em>bank_bits</em> tutorial. This tutorial has been removed.</td>
</tr>
<tr>
<td>2019.09.10</td>
<td>19.2</td>
<td>• Corrected typo in the description of the <strong>-c</strong> option in <em>Intel HLS Compiler Pro Edition Command Options</em> on page 6. The sentence that began, &quot;When you later compile the .o file...&quot; has been corrected to say, &quot;When you later link the .o file&quot;.</td>
</tr>
</tbody>
</table>

Ongoing

<table>
<thead>
<tr>
<th>Document Version</th>
<th>Intel Quartus Prime Version</th>
<th>Changes</th>
</tr>
</thead>
</table>
| 2019.07.01       | 19.2                       | • **PRO** Added information about datapath pipelining control to the following sections:  
- Loop Pipelining Control (disable_loop_pipeliningPragma) on page 68  
- Intel HLS Compiler Pro Edition Loop Pragmas on page 161  
- Component Pipelining Control (hls_disable_component_pipeliningAttribute) on page 78  
- Intel HLS Compiler Pro Edition Component Attributes on page 168  
• Revised and updated the following topics about supported math functions:  
- Math Functions Provided by the math.h Header File on page 204  
- Math Functions Provided by the extendedmath.h Header File on page 208 |
| 2019.06.04       | 19.1                       | • **PRO** In Slave Memories, clarified the use of memory attributes for slave memories.  
• In Component Memories (Memory Attributes) on page 52, clarified memory attributes support in Intel HLS Compiler Pro Edition and Intel HLS Compiler Standard Edition. |
| 2019.05.03       | 19.1                       | • **PRO** Added information about the ihc_hls_set_component_wait_cycle testbench API function to the following sections:  
- System of Tasks Simulation on page 110  
- Intel HLS Compiler Pro Edition Simulation API (Testbench Only) on page 151  
• Updated diagrams in Task Functions on page 101.  
• Updated diagram in Intel HLS Compiler Hardware Model on page 14.  
• Updated diagram in Creating Objects From RTL Code on page 117.  
• Updated diagram in Integration of an RTL Module into the HLS Pipeline on page 118. |
| 2019.04.01       | 19.1                       | • **PRO** Added information about developing your system with HLS tasks in Systems of Tasks on page 100.  
• **PRO** Added information about templated and overloaded functions in Templated and Overloaded Functions on page 18.  
• **PRO** Added information about arbitrary precision complex number (ac_complex) support to Arbitrary Precision Math Support on page 79.  
• Updated Compiler Interoperability on page 11 with details about how to use GCC and Microsoft Visual Studio to compile your component.  
• Added information about the compiler pipeline approach in Intel HLS Compiler Hardware Model on page 14.  
• In Intel HLS Compiler Pro Edition Command Options on page 6, corrected --gcc-toolchain option syntax.  
• In Intel HLS Compiler Pro Edition Command Options on page 6, updated the description of the --quartus-compile to indicate that your component is not expected to close timing when you compile your component with this option.  
• Updated the following sections with information about the --hyper-optimized-handshaking option of the i++ command:  
- Intel HLS Compiler Pro Edition Command Options on page 6  
- Intel HLS Compiler Pro Edition i++ Command-Line Arguments on page 144 |

*continued...*
Updated Loop-Carried Dependencies (ivdepPragma) on page 62 to indicate that arrays specified by the ivdep loop pragma can now be a reference to an mm_master object.

Revised and reorganized Intel High Level Synthesis Compiler Pro Edition Reference Summary on page 144.

In Declaring ac_int Data Types on page 81, revised the advice for initializing an ac_int variable to a value larger than 64 bits. To initialize this size of ac_int variable, use the bit_fill or bit_fill_hex utility functions.

2019.01.03  18.1.1


2018.12.24  18.1.1

Removed information about the "HLS/iostream" header file. The function provided by this header file is replaced by using the standard C++ iostream header and the HLS_SYNTHESIS macro.

Added description of the HLS_SYNTHESIS macro to C and C++ Libraries on page 16.

2018.12.24  18.1

Updated Slave Interfaces and Quick Reference with information about slave memory reads and writes that come from outside of the component.

Added information about conduit creation and address spaces to Avalon Memory-Mapped Host Interfaces on page 32.

2018.09.24  18.1

The Intel HLS Compiler has a new front end. For a summary of the changes introduced by this new front end, see Improved Intel HLS Compiler Front End in the Intel HLS Compiler Version 18.1 Release Notes.

The --promote-integers flag and the best_practices/integer_promotion tutorial are no longer supported in Pro Edition because integer promotion is now done by default. The flag and tutorial are still supported in Standard Edition.

Components invoked with the hls_avalon_slave_component argument must take slave or stable arguments. If the component arguments are not slave or stable arguments, compiling the component generates an error message. The description of the hls_avalon_slave_component argument in Component Invocation Interface Control Attributes and Quick Reference now reflects that requirement.

In Loops in Components on page 60, clarified the pragma statements that apply to loops must immediately precede the loop that the pragma applies to.

In Declaring ac_int Data Types on page 81, added initialization requirement for ac_int variables larger than 64 bits. You must use ac::init_array constructors to initialize ac_int variables larger than 64 bits.

In Static Variables on page 58, removed the restriction on applying memory attributes to file-scoped static variables. Both file-scoped and function-scoped static variables can have memory attributes applied to them.

2018.07.08  18.0

In Static Variables on page 58, highlighted paragraph that says that memory attributes applied to static variables work only if the static variable is declared within the component function.

In Control and Status Register (CSR) Agent on page 43, corrected a typo. The sentence "You do not need to use the hls_avalon_slave_component attribute to use the hls_avalon_slave_component attribute was corrected to say "You do not need to use the hls_avalon_slave_component attribute to use the hls_avalon_slave_register_argument attribute".
<table>
<thead>
<tr>
<th>Document Version</th>
<th>Intel Quartus Prime Version</th>
<th>Changes</th>
</tr>
</thead>
</table>
| 2018.05.07       | 18.0                        | • Starting with Intel Quartus Prime Version 18.0, the features and devices supported by the Intel HLS Compiler depend on what edition of Intel Quartus Prime you have. Intel HLS Compiler publications now use icons to indicate content and features that apply only to a specific edition as follows:  
  
  [PRO] Indicates that a feature or content applies only to the Intel HLS Compiler provided with Intel Quartus Prime Pro Edition.  
  
  [STD] Indicates that a feature or content applies only to the Intel HLS Compiler provided with Intel Quartus Prime Standard Edition.  
  
  • Corrected the code example in Intel HLS Compiler Streaming Input Interfaces Code Example. The corrected line is `int x = a.tryRead(success);` (was `int x = a.tryRead(&success);`).  
  
  • Added `<quartus_installdir>/hls/examples/tutorials/interfaces/explicit_streams_packets_empty` to list of tutorials in Table 54 on page 181 and Quick Reference.  
  
  • Added `ihc::firstSymbolInHighOrderBits` and `ihc::usesEmpty` to the list of stream interface declarations in Table 54 on page 181 and Quick Reference. Also, revised the description of the `ihc::bitsPerSymbol` declaration to include the effect of the `ihc::firstSymbolInHighOrderBits` declaration.  
  
  • Added a footnote to the `[FB519042-march MAX10[FB519042]` option in Command Options about a prerequisite required before you synthesize your component IP for Intel MAX® 10 devices.  
  
  • Added new topic AC Data Types and Native Compilers on page 90 describing use of reference AC datatype headers with the Intel HLS Compiler.  
  
  • Advanced Math Source Code Libraries on page 198 added to document Intel HLS Compiler libraries. The following Intel HLS Compiler libraries were added:  
    • Random Number Generator Library on page 198  
    • Matrix Multiplication Library on page 199 |
| 2017.12.22       | 17.1.1                      | • Updated `hls_avalon_slave_memory_argument(N)` description in Slave Memories to include the description that the parameter value `N` is the size of the memory in bytes.  
  
  • Updated Table 20 on page 86 and Table 61 on page 197 to indicate that the `ac_int` debug macros have the following restrictions:  
    • You must declare the macros in your code before you declare `#include HLS/ac_int.h`.  
    • The `ac_int` debugging tools work only for x86 emulation of your component.  
  
  • Updated `-march "<FPGA_family>"` options in Intel HLS Compiler Pro Edition Command Options on page 6 to include FPGA family options without a space.  
  
  • Revised the description of the `ihc::align` argument in `ihc::align Template Parameter` on page 193 in Quick Reference. The same information also appears in Avalon Memory-Mapped Host Interfaces on page 32.  

continued...
<table>
<thead>
<tr>
<th>Document Version</th>
<th>Intel Quartus Prime Version</th>
<th>Changes</th>
</tr>
</thead>
</table>
| 2017.11.06       | 17.1                       | • Updated *Intel HLS Compiler Pro Edition Command Options* on page 6 as follows:  
  — Revised description of `-c i++` command option.  
  — Added descriptions of the `--x86-only` and `--fpga-only i++` command options.  
• Updated *Supported Math Functions* on page 204 as follows:  
  — Noted that the `HLS/extendedmath.h` header file is supported only by the Intel HLS Compiler, not by the GCC or MSVC compilers.  
  — Added `popcount` to the list functions supported by the HLS/extendedmath.h header file.  
  — Expanded list of functions provided by HLS/extendedmath.h to explicitly list double-precision and single-precision floating point versions of the functions.  
  — Added a list of `popcount` function variations available for different data types.  
• Updated *Arbitrary Precision Math Support* on page 79 to include restriction that the Intel arbitrary precision header files cannot be compiled with GCC.  
• Added the `ihc::readwrite_mode` Avalon-MM interface to *Avalon Memory-Mapped Host Interfaces* on page 32 and *Quick Reference*.  
• Added the `ihc::waitrequest` Avalon-MM interface to *Avalon Memory-Mapped Host Interfaces* on page 32 and *Quick Reference*.  
• Added the `hls_stall_free_return` macro and `stall_free_return` attribute to *Stable Component Parameters* on page 48 and *Quick Reference*.  
• Reorganized the overall structure of the book, breaking up chapter 1 into smaller chapters and changing the order of the chapters.  
• Updated mentions of the HLS or i++ installation directory to use the Intel Quartus Prime Design Suite installation directory as the starting point.  
• Moved the following content to *Intel High Level Synthesis Compiler Best Practices Guide*:  
  — Moved "Avoid Pointer Aliasing" section to "Avoid Pointer Aliasing". |
| 2017.06.23       | —                          | • Updated *Static Variables* on page 58 to add information about static variable initialization and how to control it.  
  • Minor changes and corrections. |
| 2017.06.09       | —                          | • Revised *Declaring ac_int Data Types* on page 81 for changes in how to include `ac_int.h`.  
• Revised *Arbitrary Precision Math Support* on page 79 to clarify support for Algorithmic C datatypes.  
• Removed all mentions of `--device` compiler option. This option has been replaced by the changed function of the `--march` compiler option. See Table 3 on page 7 for details about the changed function of the `--march` compiler option.  
• Updated the generated C header file for the component `mycomp_xyz` in *Control and Status Register (CSR) Agent* on page 43.  
• Added information about structs in component interfaces to *Component Interfaces* on page 21.  
• Revised *C and C++ Libraries* on page 16 with updates to iostream behavior.  
• Added information about math functions supported by `extendedmath.h` header file to *Supported Math Functions* on page 204. |
<p>| 2017.02.03       | —                          | • In <em>Scalar Parameters and Avalon Streaming Interfaces</em>, updated information in the <em>Available Scalar Parameters for Avalon-ST Interfaces</em> table. |</p>
<table>
<thead>
<tr>
<th>Document Version</th>
<th>Intel Quartus Prime Version</th>
<th>Changes</th>
</tr>
</thead>
</table>
| 2016.11.30       |                             | • In HLS Compiler Command Options, modified the table Command Options that Customize Compilation in the following manner:  
  — Removed the \(--rtl-only\) command option and its description because it is no longer in use.  
  — Added the \(--simulator <name>\) command option and its description.  
  — Remove the \(-g\) command option because the HLS compiler now generates debug information in reports by default for both Windows and Linux. In addition, debug data is available by default in final binaries for Linux.  
• In Pointer Parameters, Reference Parameters, and Avalon Memory-Mapped Master Interfaces, added information on the `altera::align<value>` template argument in the table.  
• Added the topics Memory-Mapped Test Bench Constructor and Implicit and Explicit Examples of Creating a Memory-Mapped Master Test Bench.  
• In Usage Examples of Component Invocation Protocol Macros, replaced component invocation protocol attributes in the code examples with their corresponding macros.  
• Added the line `#include "HLS/hls.h"` to the code snippets in the following sections:  
  — Usage Examples of Interface Synthesis Macros  
  — Usage Examples of Component Invocation Protocol Macros  
• Added the topic Arbitrary Precision Integer Support to introduce the `ac_int` datatype and the Intel-provided `ac_int.h` header file.  
  Included the following subtopics:  
  — Defining the `ac_int` Datatype in Your Component for Arbitrary Precision Integer Support  
  — Important Usage Information on the `ac_int` Datatype  
• Updated the content in Area Minimization and Control of On-Chip Memory Architecture:  
  — Replaced the `numreadports(n)` and `numwriteports(n)` entries the Attributes for Controlling On-Chip Memory Architecture table with a single `numports_readonly_writeonly(m,n)` entry.  
  — Added information on the `hls_simple_dual_port_memory` macro.  
  — Added information on the `hls_merge ("label","direction")` and the `hls_bankbits(b0, b1, ..., bn)` attributes.  
• Added example use cases for the `hls_merge("label","direction")` and the `hls_bankbits(b0, b1, ..., bn)` attributes.  
• Added the topic Relationship between hls_bankbits Specifications and Memory Address Bits to explain the derivation of a memory address in the presence of the `hls_bankbits` and `hls_bankwidth` attributes. |
| 2016.09.12       |                             | Initial release. |