

# High-Level Synthesis: Bridging Software and Hardware

Romina Soledad Molina, Ph.D. MLab-STI, ICTP

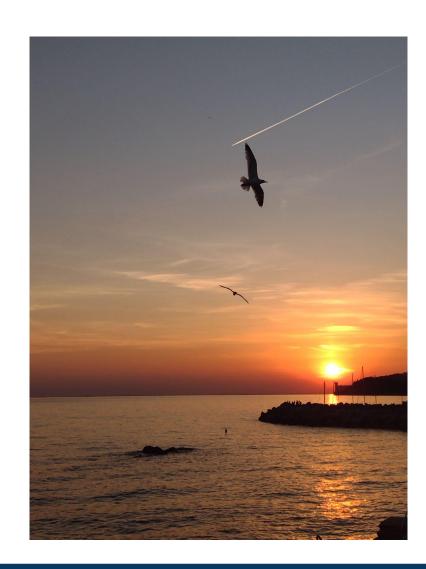
Perú - Online - 2025 -



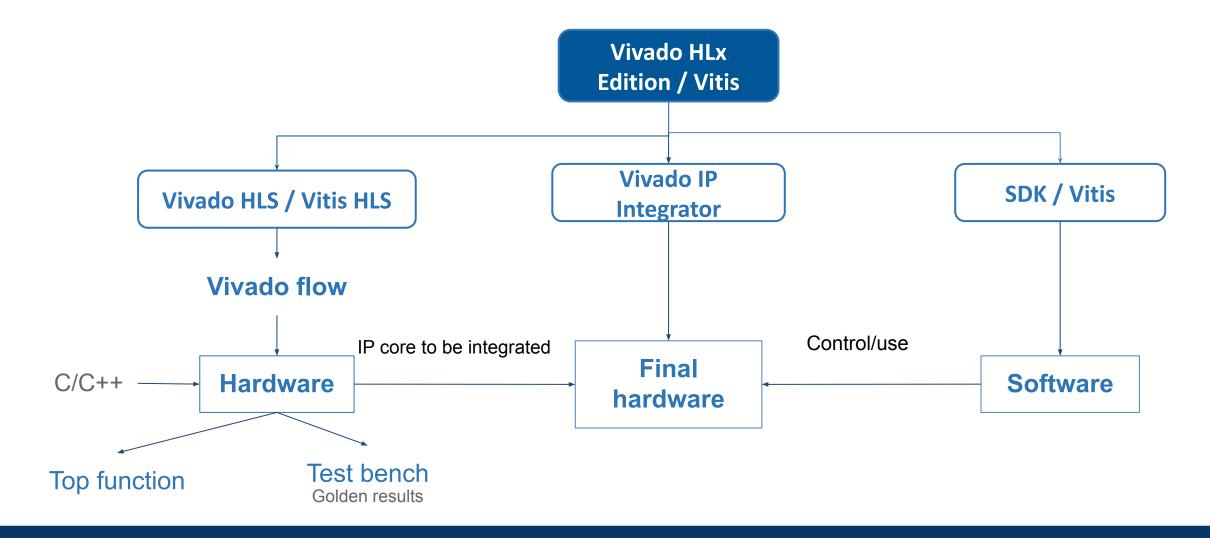


# **Outline**

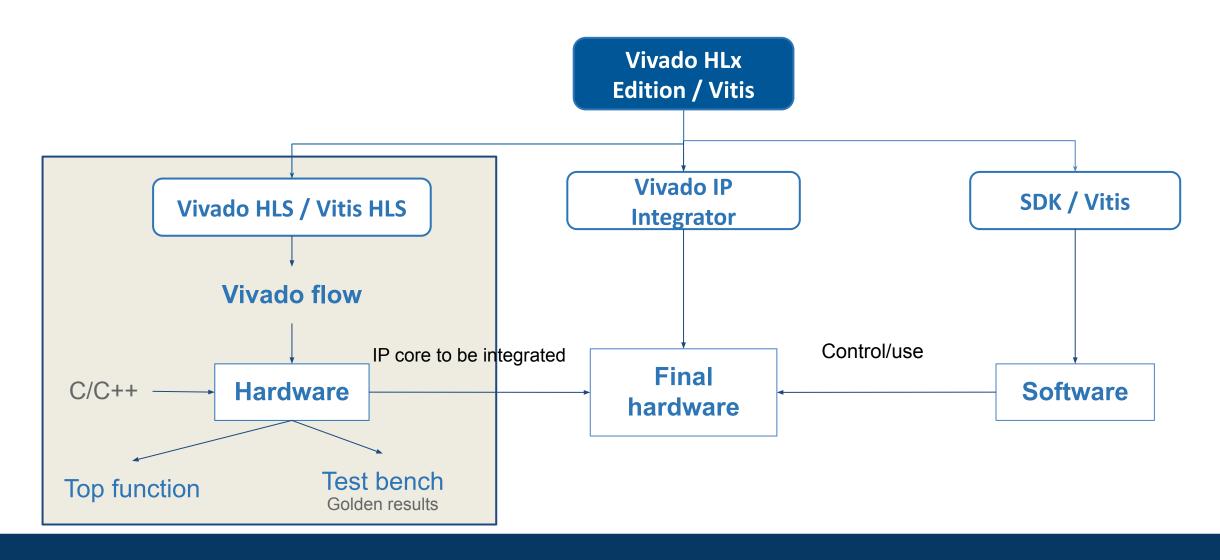
- Introduction.
- High-level synthesis.
- HLS Component Development Flow.
- C-to-RTL Conversion.
- Language Support.
- Hardware design: Directives/Optimizations.
- Vitis HLS GUI.
- Demo: HLS and Matrix Multiplication.





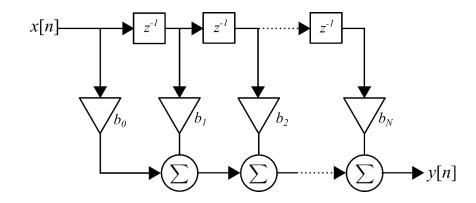


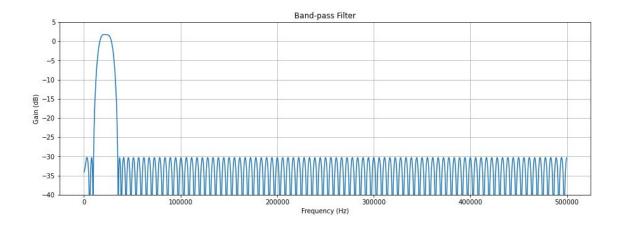






- Traditional RTL FIR Filter Design
  - Define interfaces
  - Define architecture
    - FSM
    - Datapath
  - Write RTL code
  - Write RTL test bench.
- The design choice is already made.







• Question: Why the need of High-Level Synthesis tools?



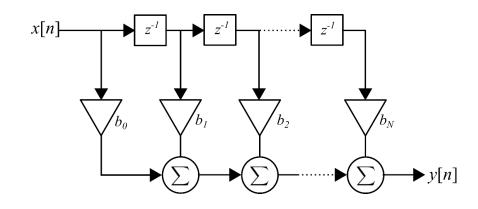


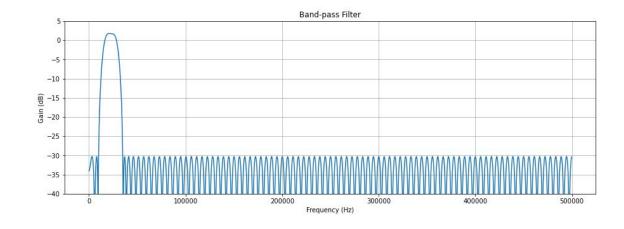
- Question: Why the need of High-Level Synthesis tools?
  - Some reasons could be:
    - Productivity boosting.
    - Trend to use FPGA as hardware accelerators.
    - Reduce Time-To-Market.
    - Design Space Exploration.
    - Early metric estimations.
    - Functionality verification through C-based test bench.
    - Reuse.



#### HLS-based FIR Filter Design

```
void fir (data_t *y, data_t *x ) {
      static data_t shift_reg[N];
      acc_t acc;
      data_t data;
      int i;
      acc=0;
      Shift_Accum_Loop: for (i=N-1;i>=0;i--) {
             if (i==0) {
                   shift_reg[0] = *x;
                   data = *x;
             } else {
                   shift_reg[i]=shift_reg[i-1];
                   data = shift_reg[i];
      acc+=data*firCoeff[i];
      acc = acc >> 16;
      *y=acc;
```





# **High-Level Synthesis**



- High-Level Synthesis
  - It provides the facility to create RTL from a high level of abstraction.



- High-Level Synthesis
  - It provides the facility to create RTL from a high level of abstraction.
  - Several implementations are possible from the same source code description.



#### High-Level Synthesis

- It provides the facility to create RTL from a high level of abstraction.
- Several implementations are possible from the same source code description.
- Implements the design based on defaults and user applied directives.



#### High-Level Synthesis

- It provides the facility to create RTL from a high level of abstraction.
- Several implementations are possible from the same source code description.
- Implements the design based on defaults and user applied directives.
- It allows the optimization of the input code using directives to:
  - Reduce latency
  - Improve performance and throughput
  - Reduce resource utilization

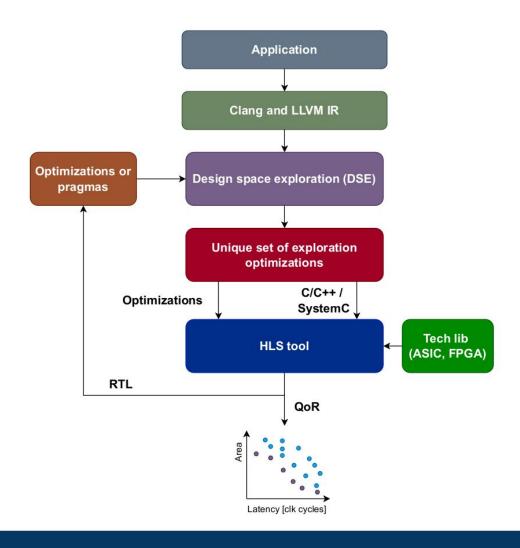


#### High-Level Synthesis

- It provides the facility to create RTL from a high level of abstraction.
- Several implementations are possible from the same source code description.
- Implements the design based on defaults and user applied directives.
- It allows the optimization of the input code using directives to:
  - Reduce latency.
  - Improve performance and throughput.
  - Reduce resource utilization.

Without optimization, HLS tool will look to minimize latency and improve concurrency.

# **Design space exploration**



# **Design space exploration**

```
...
loop: for (i=3;i>=0;i--) {
    if (i==0) {
        acc+=x*c[0];
        shift_reg[0]=x;
    } else {
        shift_reg[i]=shift_reg[i-1];
        acc+=shift_reg[i]*c[i];
}
```

Same hardware is used for each loop iteration:

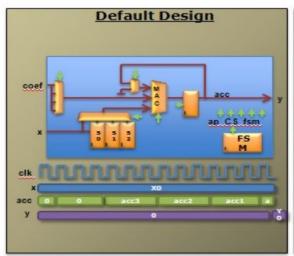
- Small area
- Long latency
- Low throughput

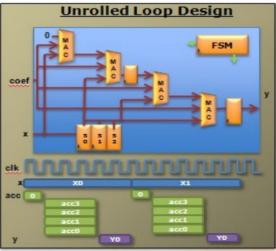
Different hardware for each loop iteration:

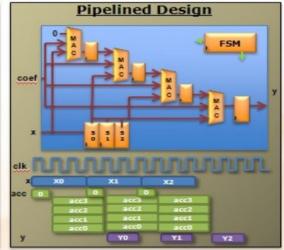
- Higher area
- Short latency
- Better throughput

Different iterations executed concurrently:

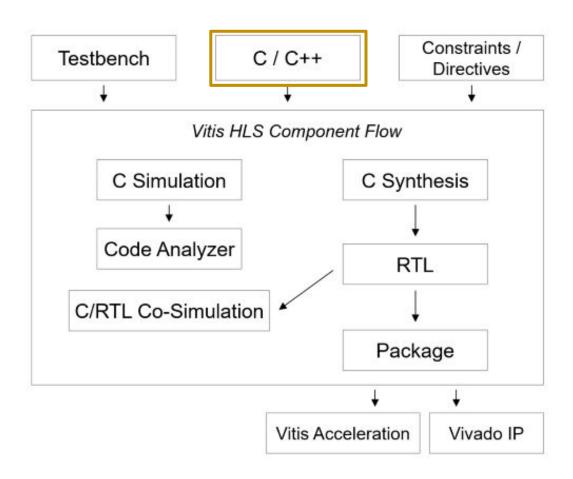
- Higher area
- · Short latency
- Best throughput



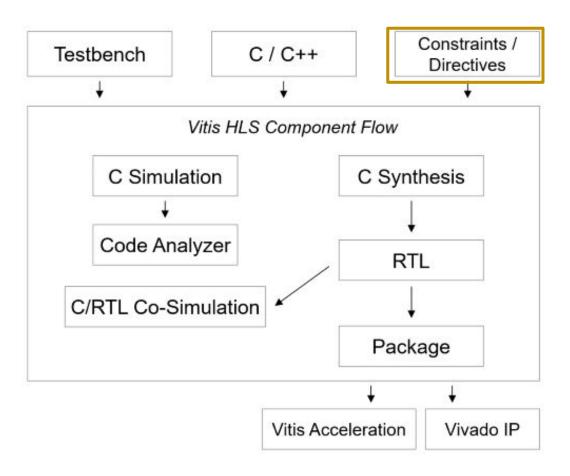




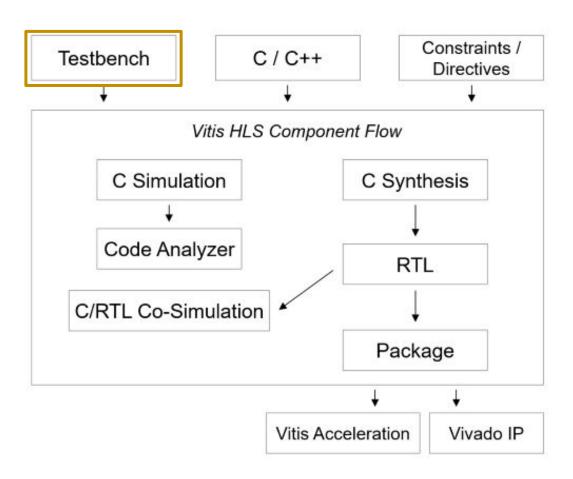




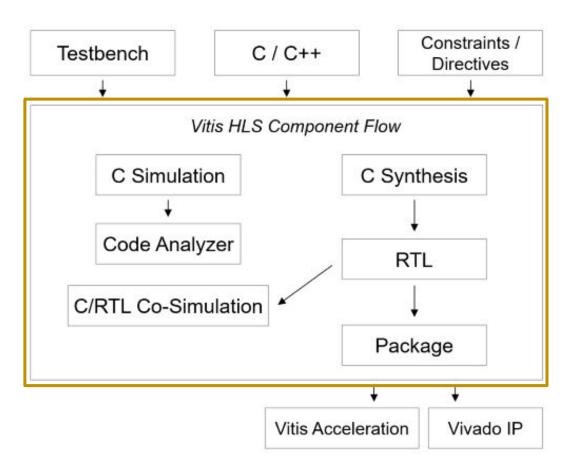




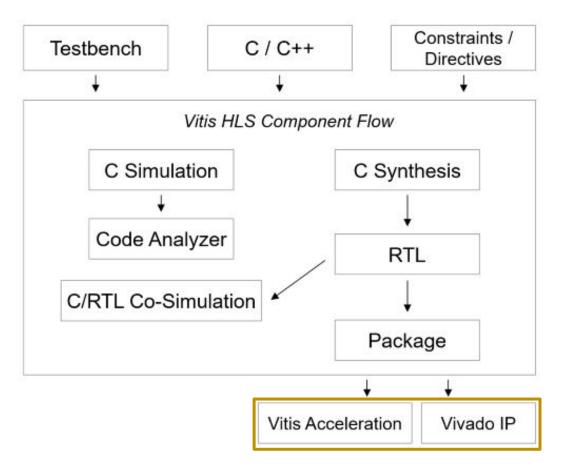








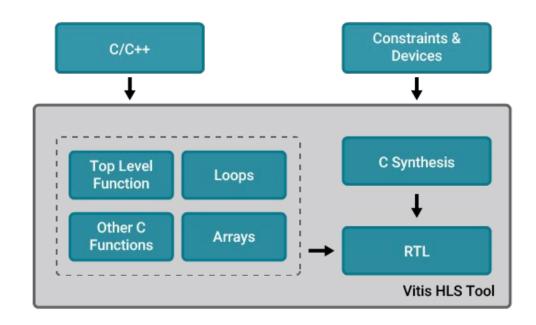






The Vitis HLS tool synthesizes different parts of C code differently:

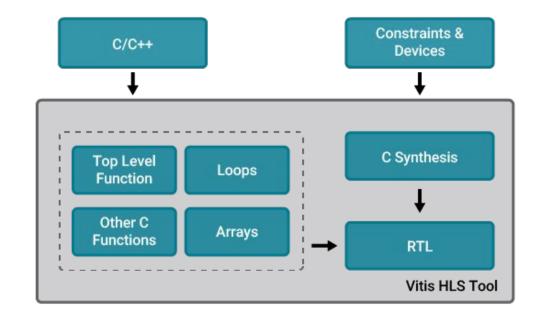
 Top-level function arguments of the C/C++ code are synthesized into RTL I/O ports and are automatically implemented with an interface synthesis hardware protocol -> Only one top function.





The Vitis HLS tool synthesizes different parts of C code differently:

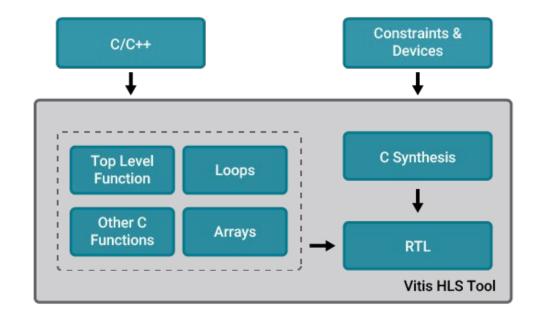
 Top-level function arguments of the C/C++ code are synthesized into RTL I/O ports and are automatically implemented with an interface synthesis hardware protocol → Only one top function.





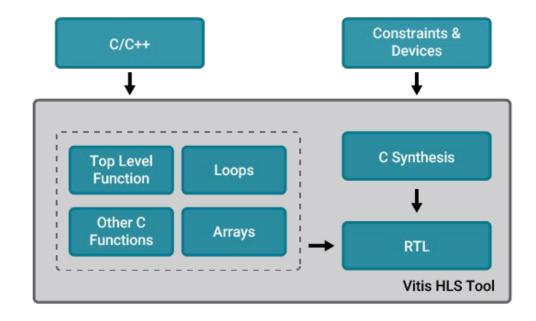
The Vitis HLS tool synthesizes different parts of C code differently:

- Top-level function arguments of the C/C++ code are synthesized into RTL I/O ports and are automatically implemented with an interface synthesis hardware protocol → Only one top function.
- Other C functions are synthesized to RTL blocks—maintaining the design hierarchy.
- Arrays in the C code can be targeted to any memory resource, such as BRAM, LUTRAM, and URAM.





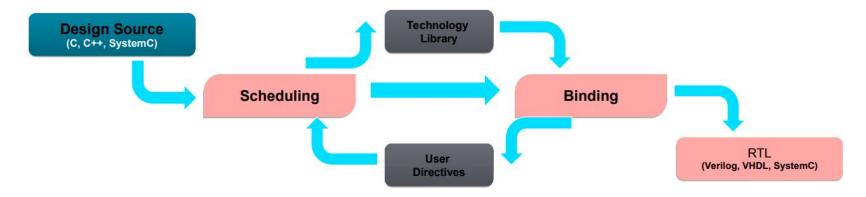
- Performance metrics, such as latency, initiation interval, loop iteration latency, and resource utilization, can be reviewed with synthesis reports.
- Vitis HLS tool **pragmas and optimization directives** allow for configuring the synthesis results for the C/C++ code.





# **C-to-RTL Conversion**Scheduling and binding

- Scheduling determines in which clock cycle an operation will occur.
- **Binding** determines which library cell is used for each operation.
  - For example, for a functional unit like adder, there can be many options like ripple-carry adder, carry-look-ahead-adder etc.





# **C-to-RTL Conversion**Scheduling and binding

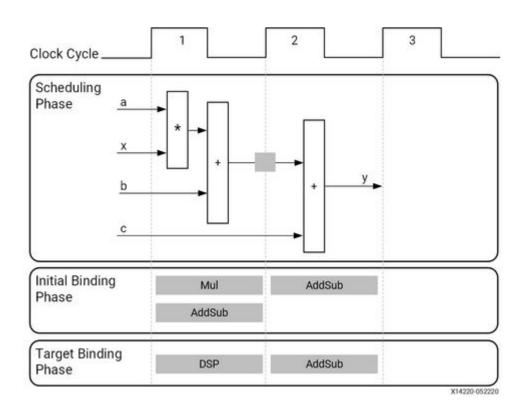
```
int foo(char x, char a, char b, char c) {
    char y;
    y = x*a+b+c;
    return y;
}
```

Source: https://docs.amd.com/r/2020.2-English/ug1399-vitis-hls/Scheduling-and-Binding-Example



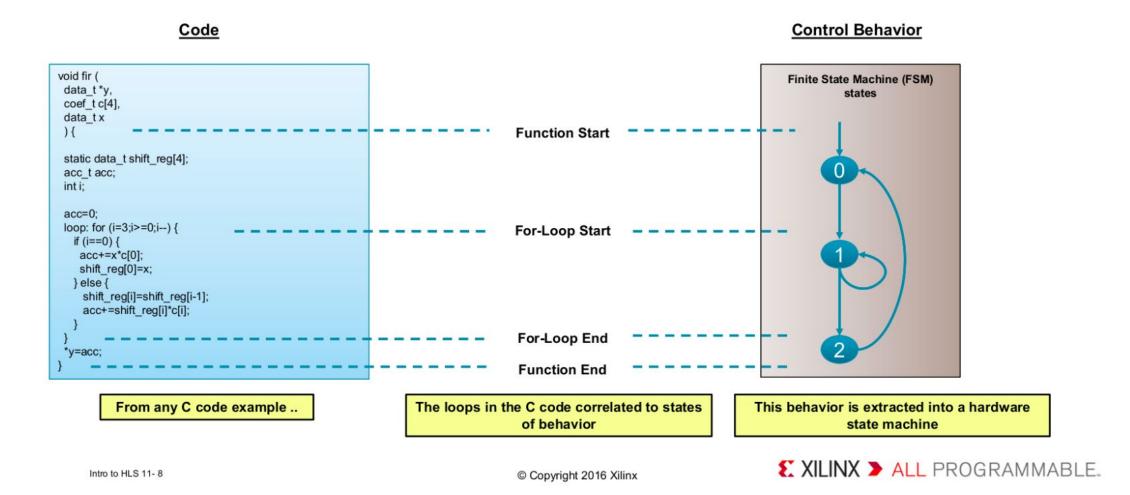
# **C-to-RTL Conversion**Scheduling and binding

```
int foo(char x, char a, char b, char c) {
    char y;
    y = x*a+b+c;
    return y;
}
```

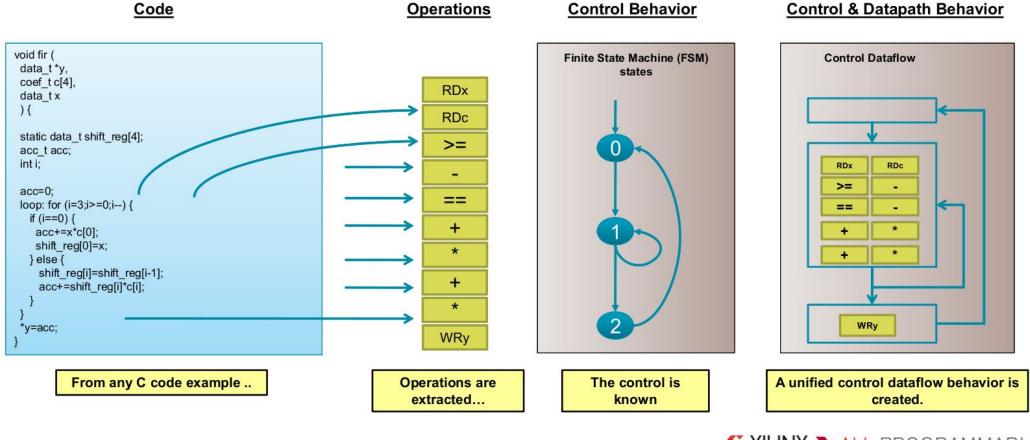


Source: https://docs.amd.com/r/2020.2-English/ug1399-vitis-hls/Scheduling-and-Binding-Example









# **Language Support**



# **Language Support**

- Vivado/Vitis HLS supports C, C++, SystemC, and OpenCL API C kernel.
- Supports arbitrary precision types for all input languages.
- Floating point support.
- Support for OpenCV functions.



# **Language Support**

- The function must contain the entire functionality of the design.
- None of the functionality can be performed by system calls to the operating system.
- The C/C++ constructs must be of a fixed or bounded size.
- The implementation of those constructs must be unambiguous.

Source: https://docs.amd.com/r/2020.2-English/ug1399-vitis-hls/Vitis-HLS-Coding-Styles



### **Language Support**

- The function must contain the entire functionality of the design.
- None of the functionality can be performed by system calls to the operating system.
- The C/C++ constructs must be of a fixed or bounded size.
- The implementation of those constructs must be unambiguous.
- Unsupported C/C++ Constructs:
  - System calls.
  - Dynamic memory usage.
  - Recursive functions.

Source: https://docs.amd.com/r/2020.2-English/ug1399-vitis-hls/Vitis-HLS-Coding-Styles

### Data type precision

- Standard C Types
  - Integers:
    - long long (64 bits)
    - int (32 bits)
    - Short (16 bits)
  - Characters:
    - char (8 bits)
  - Floating Point:
    - Float (32 bits)
    - Double (64 bits)

- Arbitrary Precision Types
  - o C:
    - ap\_(u)int
  - o C++:
    - ap\_(u)int
    - ap\_fixed
  - O C++ / SystemC:
    - sc\_(u)int
    - sc\_fixed

# Hardware design: Directives/Optimizations



## **Hardware design - Directives/Optimizations**

Minimize latency: UNROLL, LOOP\_FLATTEN, LOOP\_MERGE.

Minimize throughput: DATAFLOW, PIPELINE.

Improve bottleneck: RESOURCE, ARRAY\_PARTITION, ARRAY\_RESHAPE.

### **Hardware design - Directives/Optimizations**

Minimize latency: UNROLL, LOOP\_FLATTEN, LOOP\_MERGE.

Minimize throughput: DATAFLOW, PIPELINE.

Improve bottleneck: RESOURCE, ARRAY\_PARTITION, ARRAY\_RESHAPE.

```
#pragma HLS UNROLL
#pragma HLS PIPELINE
#pragma HLS ARRAY_PARTITION variable=layer3_out complete dim=0
```



## **Hardware design - Directives/Optimizations**

Minimize latency: UNROLL, LOOP\_FLATTEN, LOOP\_MERGE.

Minimize throughput: DATAFLOW, PIPELINE.

Improve bottleneck: RESOURCE, ARRAY\_PARTITION, ARRAY\_RESHAPE.

```
#pragma HLS UNROLL
#pragma HLS PIPELINE
#pragma HLS ARRAY_PARTITION variable=layer3_out complete dim=0
```

Source code: directives are included in the code.

Directive file: directives are specified in a separated file.

OP RD

OP CMP

OP WR

```
Clock (ckl)
Loop 1: for(i=1; i<3; i++){
   OP_RD;
    OP_CMP;
   OP_WR;
                                           OP RD
                                                        OP CMP
                                                                      OP WR
   OP RD
                 Read
                                           Initiation interval = 3 clock cycles
  OP CMP
                 Computation
                                           Latency = 3 clock cycles
   OP WR
                 Write
                                           Loop latency = 6 clock cycles
```

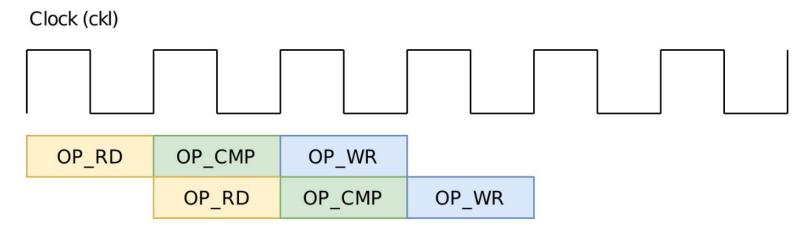
#### **Loop + Pipeline**

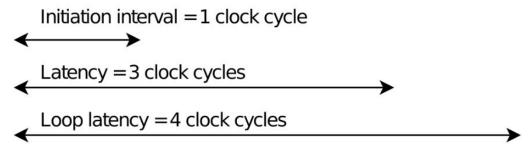
```
Loop_1: for(i=1; i<3; i++){
    OP_RD;
    OP_CMP;
    OP_WR;
}
```

OP RD Read

OP\_CMP Computation

OP\_WR Write





Latency = 3 clock cycles

Loop latency = 3 clock cycles

#### Loop + Unroll

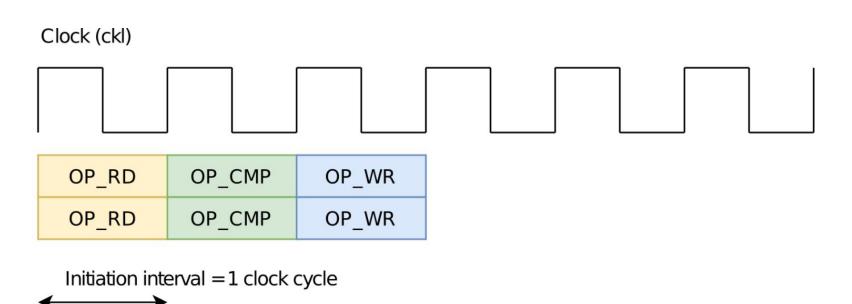
OP\_WR

```
Loop_1: for(i=1; i<3; i++){
    OP_RD;
    OP_CMP;
    OP_WR;
}

OP_RD Read

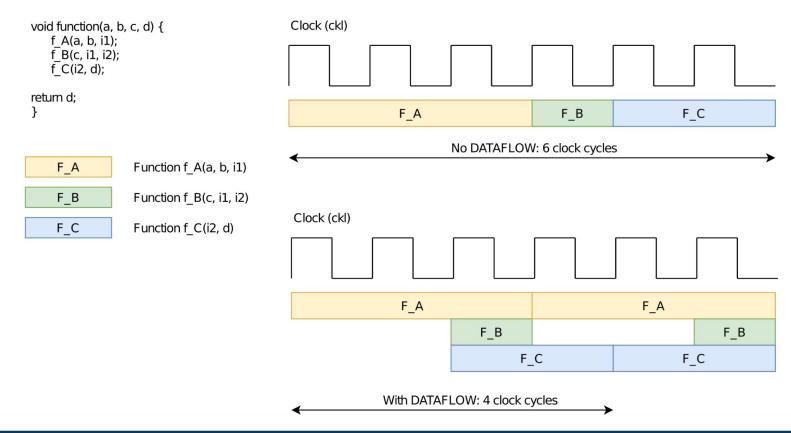
OP_CMP Computation
```

Write

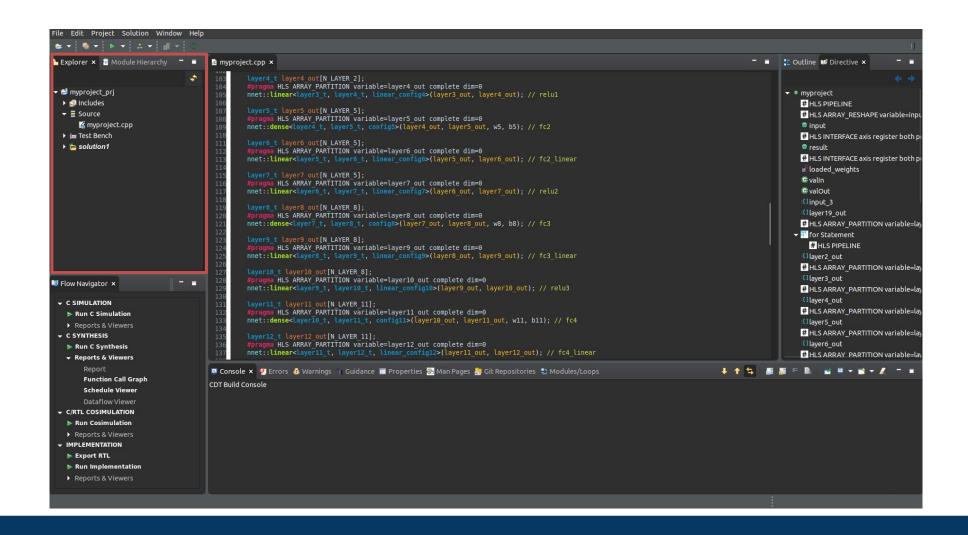




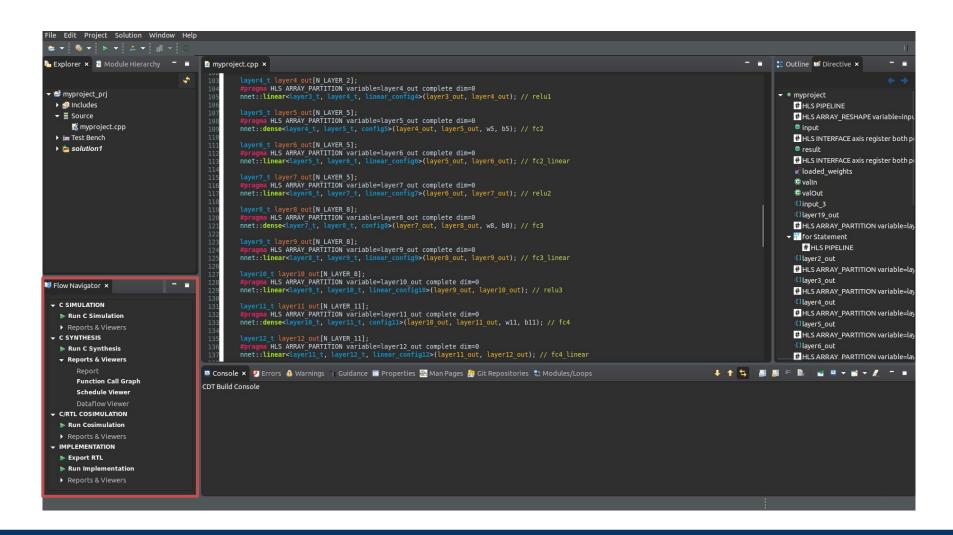
#### **Loop + Dataflow**



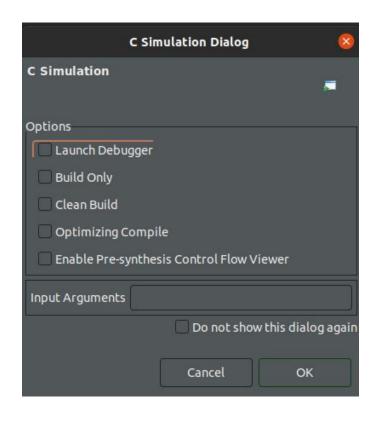


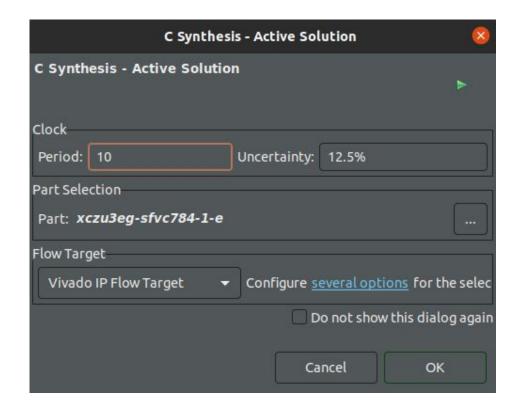










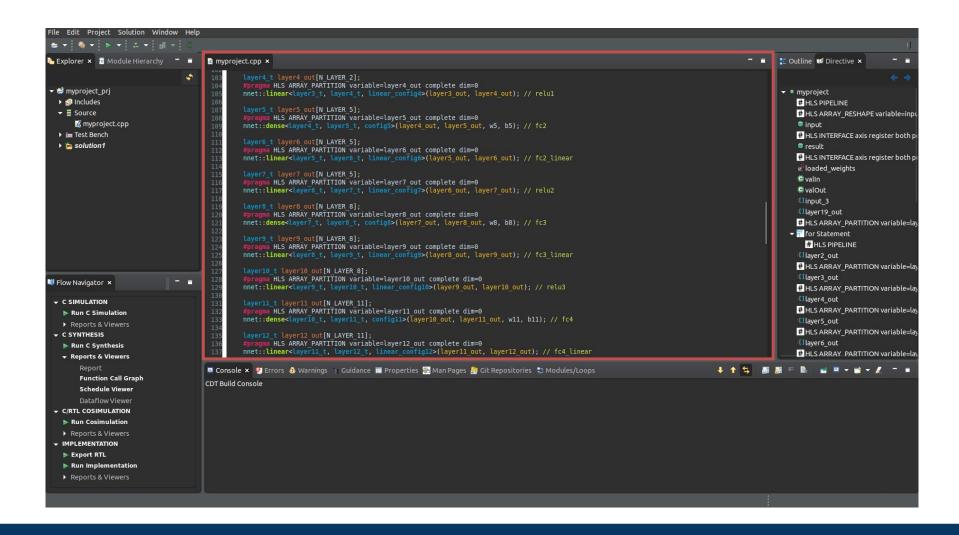




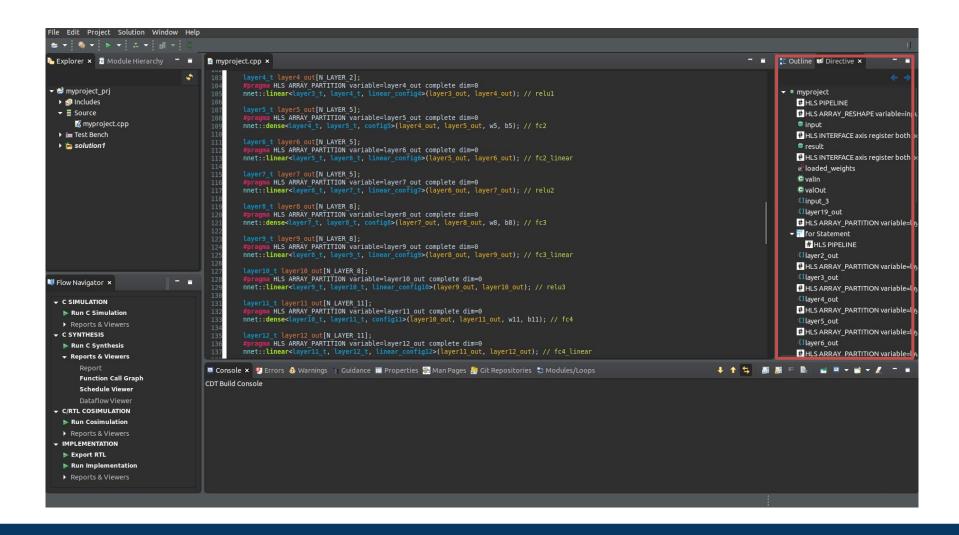
Co-simulation Dialog			
C/RTL Co-simulation ☑			
RTL Simulator Settings			
Vivado XSIM   ◆ Verilog ○ VHDL			
Setup Only			
Optimizing Compile			
Input Arguments			
Dump Trace all ▼			
Random Stall			
Compiled Library Location Browse			
Extra Options for DATAFLOW			
☐ Wave Debug (Vivado XSIM only and "Dump Trace" != none)			
☐ Disable Deadlock Detection			
Channel (PIPO/FIFO) Profiling			
Dynamic Deadlock Prevention			
☐ Do not show this dialog again			
Cancel OK			

	Export RTL	8
Export RTL as II	Р/ХО	=
Export Format	Vivado IP (.zip) ▼	
Output Location		Browse
IP OOC XDC File		Browse
IP XDC File		Browse
IP Configuration		
Vendor		
Library		
Version		
Description		
	ammaNeutron_smr3983	
Taxonomy		
	Do not show this	dialog again
	Do not show this	diatog again
	Cancel	ОК

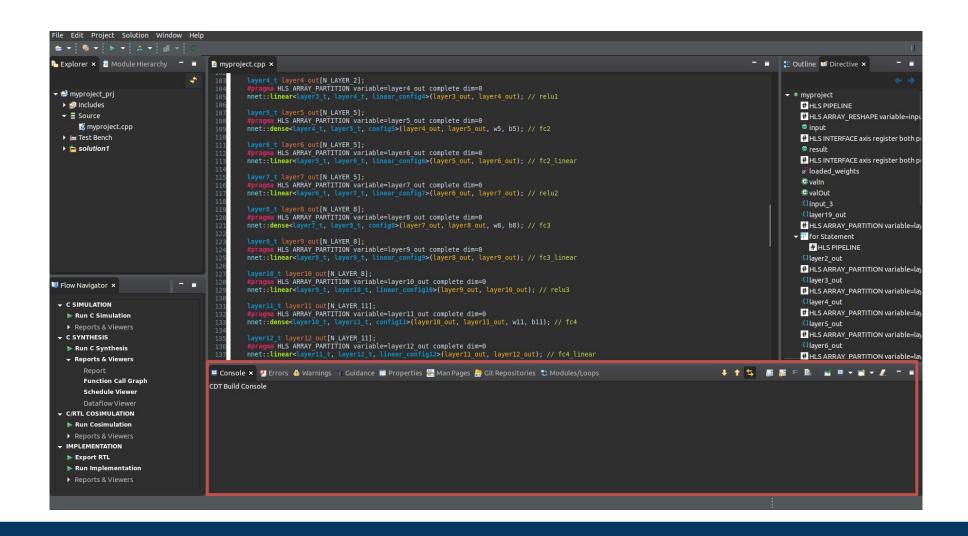






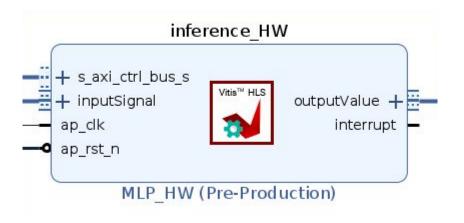






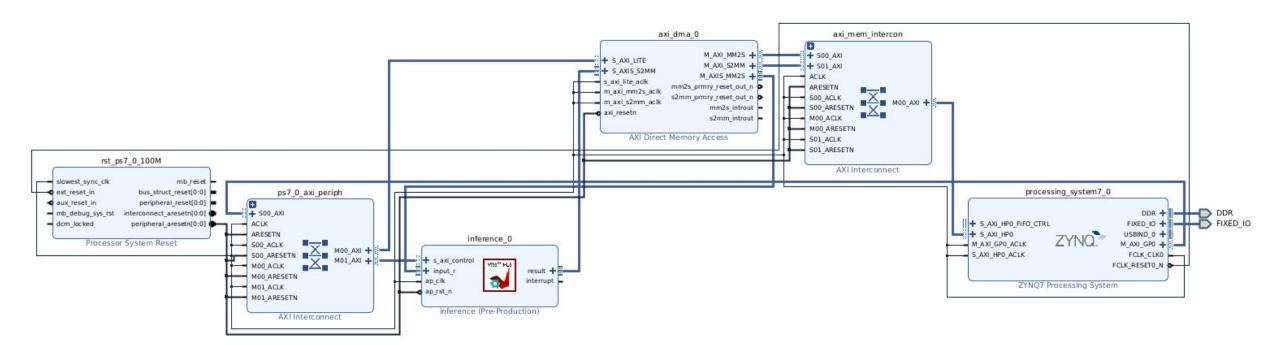


#### Generated IP core

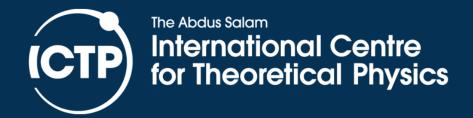


## **Vivado IP integration**

## **Vivado IP integration**



## Demo HLS: Vector Addition and Matrix Multiplication



## High-Level Synthesis: Bridging Software and Hardware

Romina Soledad Molina, Ph.D. MLab-STI, ICTP

Perú - Online - 2025 -

