



What is Vitis?

Xilinx unified software development environment

 Enables development of embedded, accelerated and AI solutions on Xilinx devices

Works hand in hand with Vivado and Petalinux

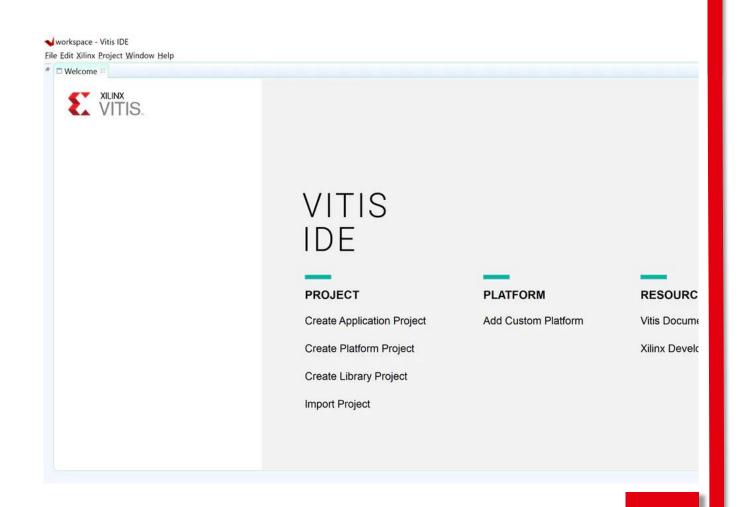
- Works on Linux and Win 10
 - Linux only supported for acceleration flow



Key concepts of Vitis

Slightly different concepts from SDK

- Platform Project
 - HW & SW resources e.g. BSP, FSBL, HW Spec
 - Grouped as domains
 - BSP, OS etc.
- System Project
 - Actual applications which are created
- Debugging on target and on QEMU
 - System Debugger





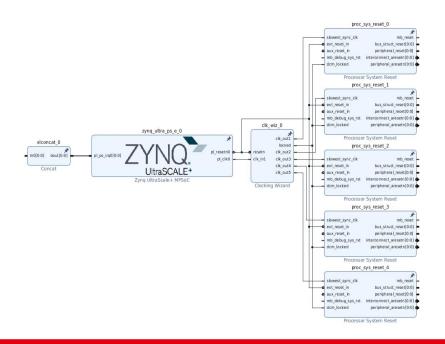
Vitis Flows

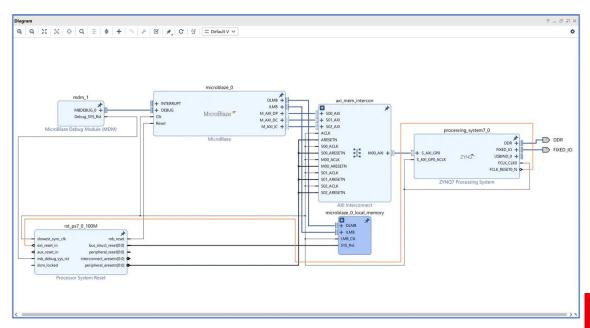
Accelerated

- Leverages OpenCL to accelerate functions into the programmable logic
- Uses Xilinx Run Time (XRT)
- Petalinux Only
- Requires specific configuration in Vivado and Petalinux

Embedded

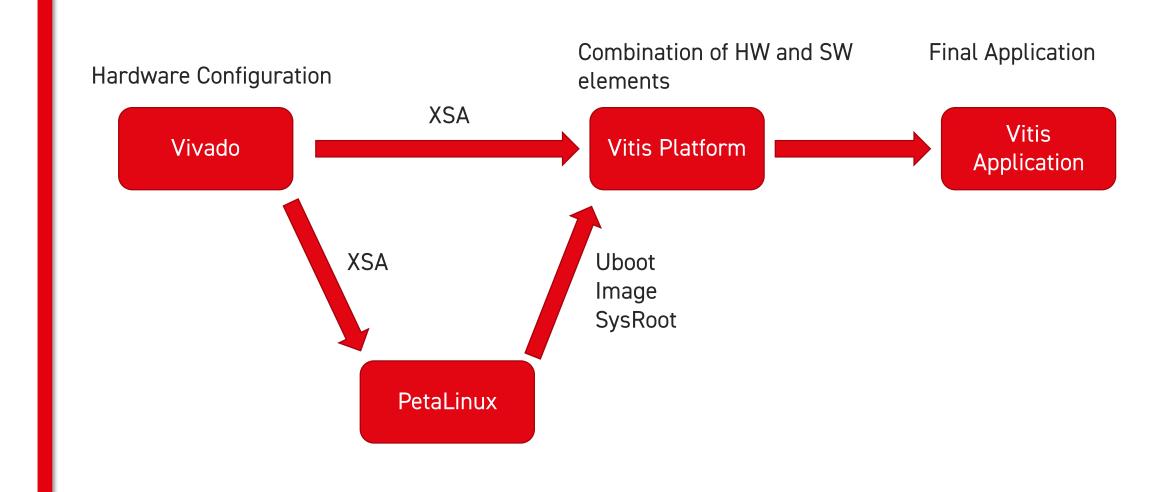
- Traditional embedded development flow
 - Bare Metal
 - Free RTOS
 - PetaLinux
- Supports A9, A53, R5, PMU & Microblaze







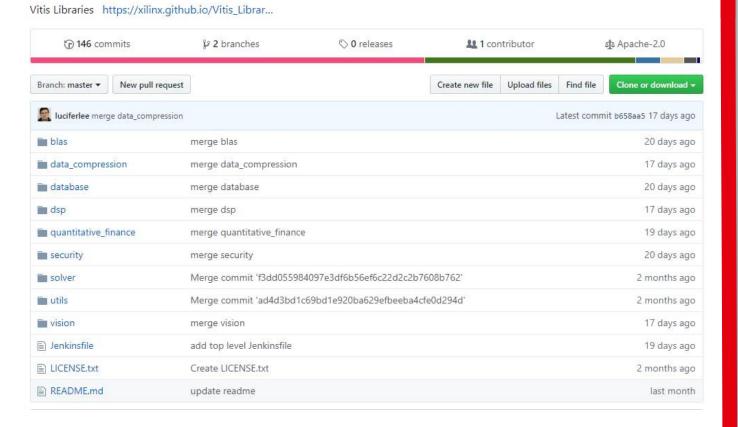
Vitis Flow





Vitis libraries

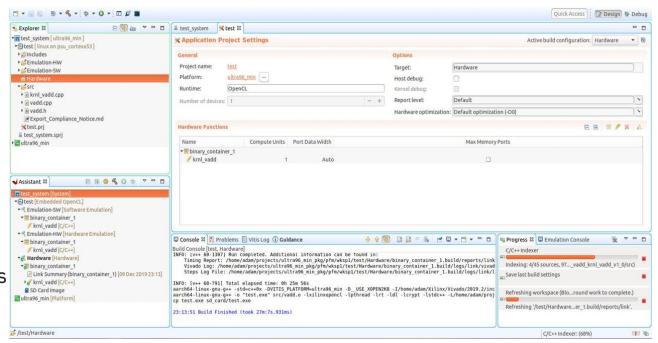
- Common libraries: These libraries provide basic functions which are used across a range of applications and domains for example math, DSP and linear algebra.
- Domain specific: These libraries provide acceleration functions for specific domains, e.g. security, vision, finance or database.
- Level one: The lowest level of implementation, intended for use in a High-Level Synthesis flow.
- Level two: Middle level provides acceleration kernels that are used in the Vitis design flow and the Xilinx RunTime (XRT).
- Level three: Highest level provides applications created from several acceleration kernels. This uses API and of course XRT.





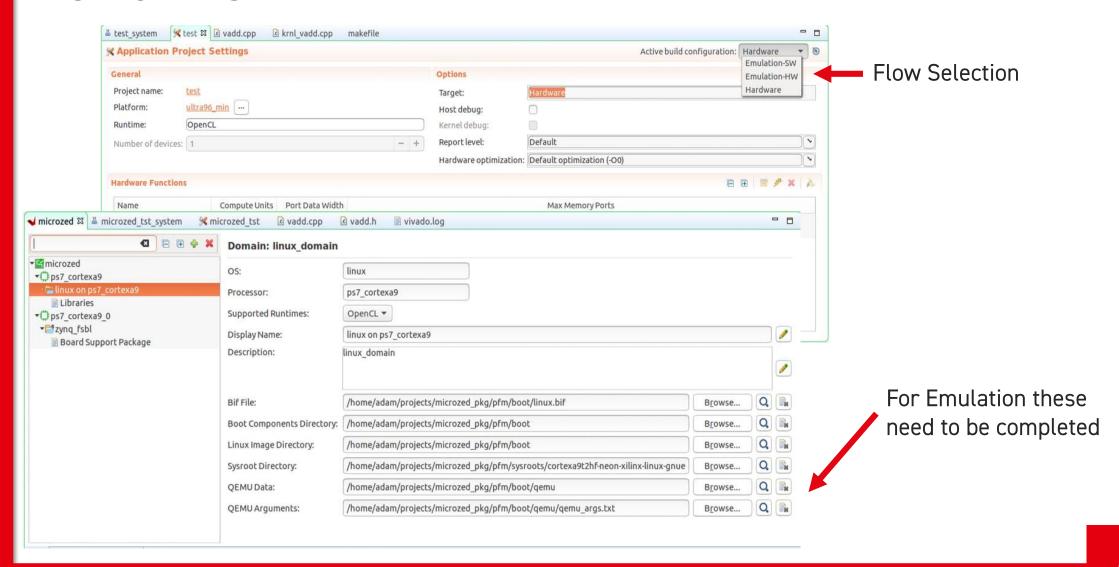
Build Flow

- Accelerated flow, building the design each time is time consuming
- Three different build flows are available
 - Software emulation: Runs both the host and the kernel on the x86.
 - Hardware emulation: Runs the kernel as a compiled hardware model, while the host uses a C simulator.
 - Hardware: Generates the final bootable image for the system





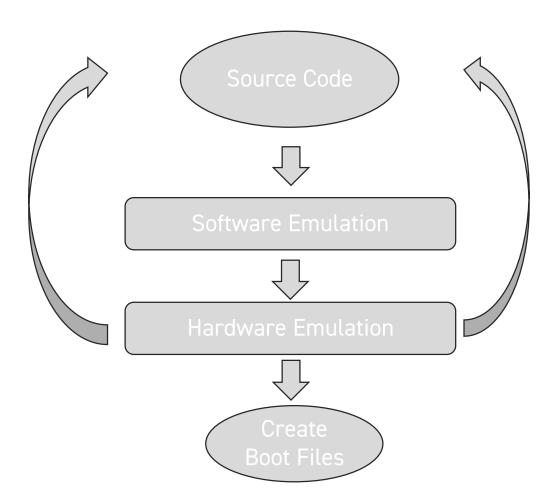
Build Flow





Emulation Flow

Optimize Performance Optimize Interfacing Optimize Resources



Syntax Errors Verify Functional Performance



Creating a Platform

- First, we need a Linux Machine
- I use a Linux VM on a USB C SSD and works very well
- AWS Instances work brilliantly too

Component	Requirement
Operating System	Linux, 64-bit: • Ubuntu 16.04.5 LTS, 16.04.6 LTS, 18.04.1 LTS, 18.04.2 LTS • CentOS 7.4, 7.5, 7.6 • RHEL 7.4, 7.5, 7.6
System Memory	For Alveo cards: 64 GB (80 GB is recommended) For embedded: 32 GB
Internet Connection	Required for downloading drivers and utilities.
Hard disk space	100 GB



XRT Run Time

- Once Vitis is installed we need to install the XRT
- Answer Record 73055 very useful if issues

```
🧶 🖱 📵 adam@adam-VirtualBox: ~/Downloads
First Installation: checking all kernels...
Building only for 4.15.0-70-generic
Building initial module for 4.15.0-70-generic
 unning module version sanity check.
  Original module
   - No original module exists within this kernel
    - Installing to /lib/modules/4.15.0-70-generic/updates/dkms/
Running module version sanity check.
   Original module
   - No original module exists within this kernel
   - Installing to /lib/modules/4.15.0-70-generic/updates/dkms/
DKMS: install completed.
Finished DKMS common.postinst
Loading new XRT Linux kernel modules
Installing MSD / MPD daemons
Skipping pyopencl installation...
 adam@adam-VirtualBox:~/Downloads$
```

Installing Xilinx Runtime @

Xilinx Runtime (XRT) is implemented as a combination of user-space and kernel driver components. XRT supports Alveo PCIe-based cards, as well as Zynq UltraScale+ MPSoC-based embedded system platforms, and provides a software interface to Xilinx programmable logic devices.

You only need to install XRT once, regardless of how many platforms you may be installing.

<u>MPORTANT: XRT installation uses standard Linux RPM and Linux DEB distribution files, and root access is required for all software and firmware installations.</u>

<rpm-dir> or <deb-dir> is the directory where you downloaded the packages to install.

To download and install the XRT package for your operating system, do the following.

∨ CentOS/RedHat ∂

- 1. To download the RPM file, click this link.
- 2. To install the package, enter the following command.

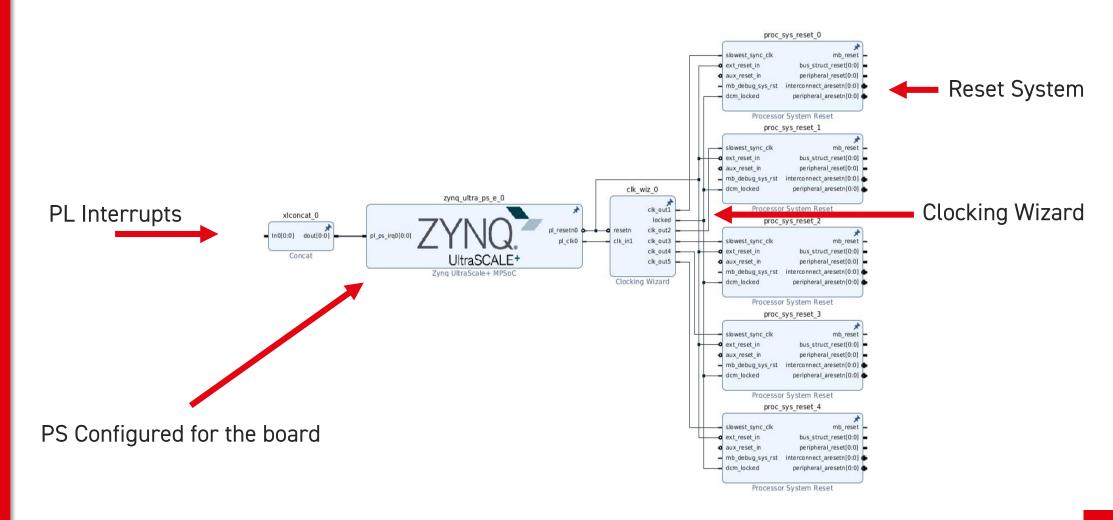
```
sudo yum install <rpm-dir>/<xrt_filename>.rpm
```

∨ Ubuntu ∂

- 1. To download the DEB file, click one of the following:
 - Ubuntu 16.04

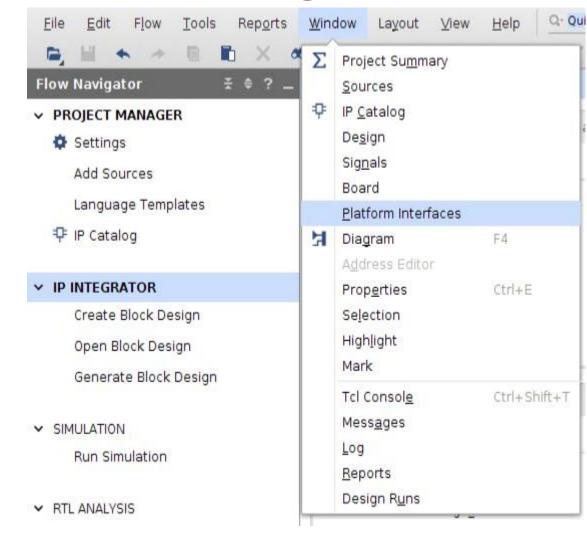


Vivado Design





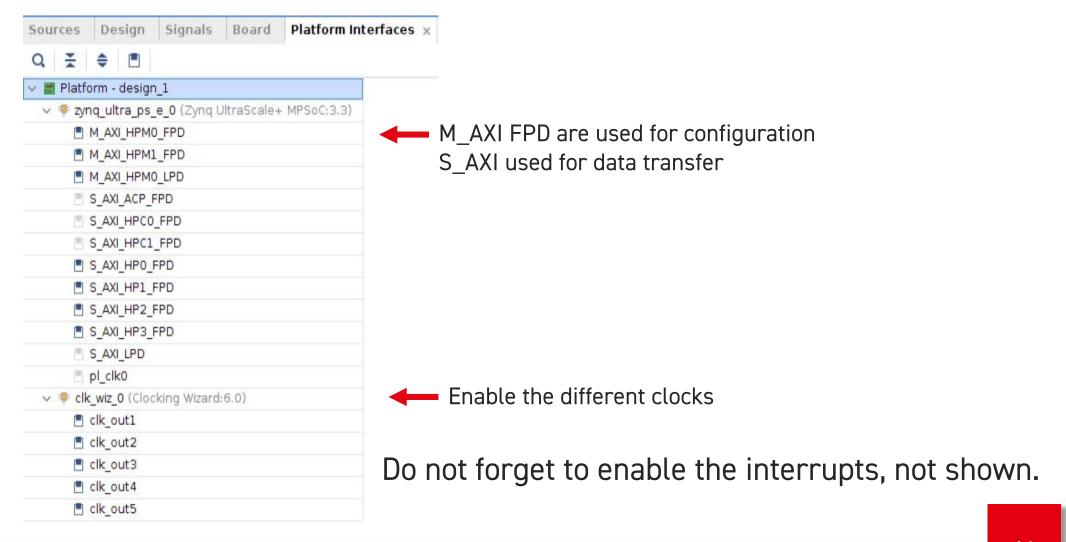
Vivado Design



Declare capabilities which may or may not be made available to the V++ compiler.



Vivado Platform Interfaces





Vivado Generating the XSA

```
set property platform.design intent.embedded true [current project]
set property platform.design intent.server managed false [current project]
                                                                                                                             Add Information to our Platform
set property platform.design intent.external host false [current project]
set property platform.design intent.datacenter false [current project]
set property platform.default output type "sd card" [current project]
  write hw platform -include bit ultra96 min.xsa
                                                                                                                                      Write out and Validate the XSA
  validate hw platform ./ultra96 min.xsa
Tcl Console x Messages Log Reports Design Runs
  INFO: [Vivado 12-4896] Successfully created Hardware Platform: /home/adam/projects/ultra96_min_pkg/vivado/ultra96_min.xsa
  write_hw_platform: Time (s): cpu = 00:00:10 ; elapsed = 00:00:10 . Memory (MB): peak = 8615.188 ; gain = 0.000 ; free physical = 10089 ; free virtual = 13700
  /home/adam/projects/ultra96_min_pkg/vivado/ultra96_min.xsa
  validate hw platform ./ultra96 min.xsa
  INFO: [Vivado 12-6074] Validating Hardware Platform: './ultra96_min.xsa'
  INFO: [Vivado 12-8115] Found XML metadata file: xsa.xml
  INFO: [Vivado 12-6078] Validating platform properties...
                                                                                                                              Validated XSA which can be exported
  INFO: [Vivado 12-6079] Validating unified platform...
  INFO: [Vivado 12-6073] Validating 'pre_synth' platform state...
  INFO: [Vivado 12-6077] Validating platform files...
  INFO: [Vivado 12-6067] Found file 'emu/emu.xml' of type 'EMU_XML' in the Hardware Platform.
  INFO: [Vivado 12-6067] Found file 'ultra96 min.bit' of type 'FULL BIT' in the Hardware Platform.
  INFO: [Vivado 12-6067] Found file 'ultra96 min.hpfm' of type 'HPFM' in the Hardware Platform.
  INFO: [Vivado 12-6067] Found file 'prj/rebuild.tcl' of type 'REBUILD_TCL' in the Hardware Platform.

☐ INFO: [Vivado 12-6066] Finished running validate dsa for file: './ultra96 min.xsa'
```



Create a new PetaLinux project based off the XSA

Need to make several modifications to the OS, Device Tree and RootFS to support XRT

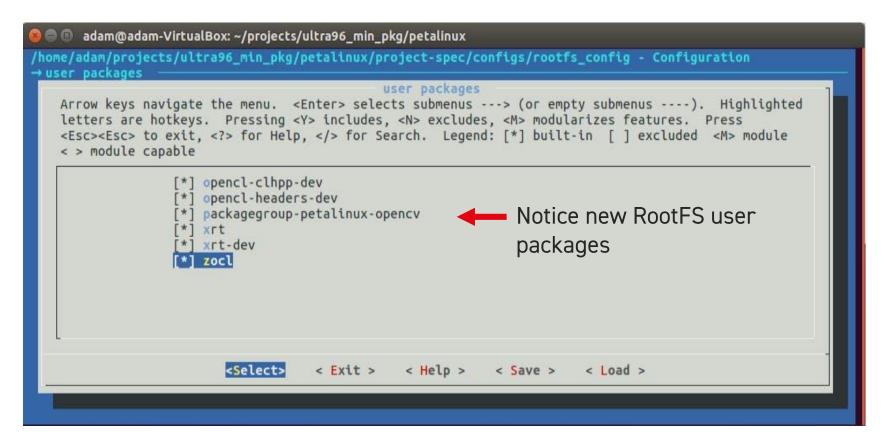








petalinux-config -c rootfs



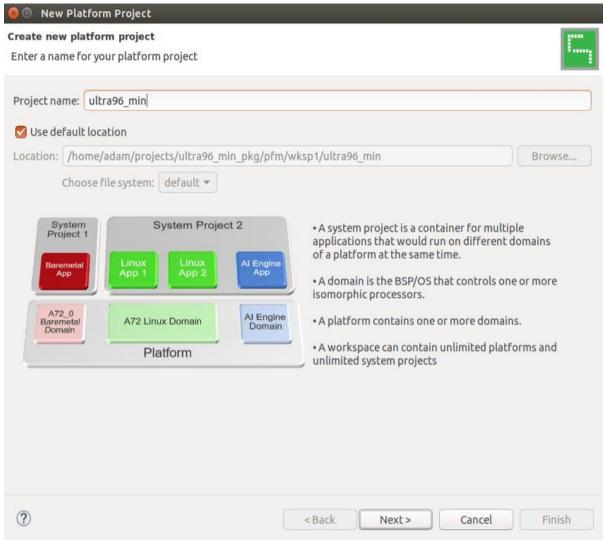


petalinux-config -c kernel

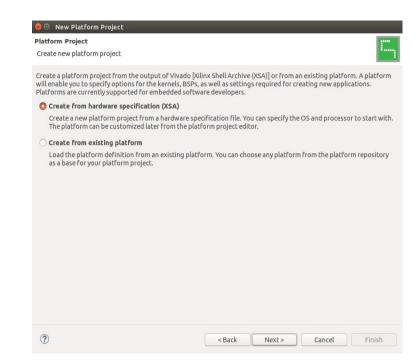
- Device Drivers > Generic Driver Options > DMA Contiguous Memory Allocator > Size in Mega Bytes change the size from 256 to 1024 MB
- Device Drivers -> Staging drivers -> Xilinx APF Accelerator driver
- Device Drivers -> Staging drivers -> Xilinx APF Accelerator driver -> Xilinx APF DMA engines support

./sdk.sh Install SysRoot

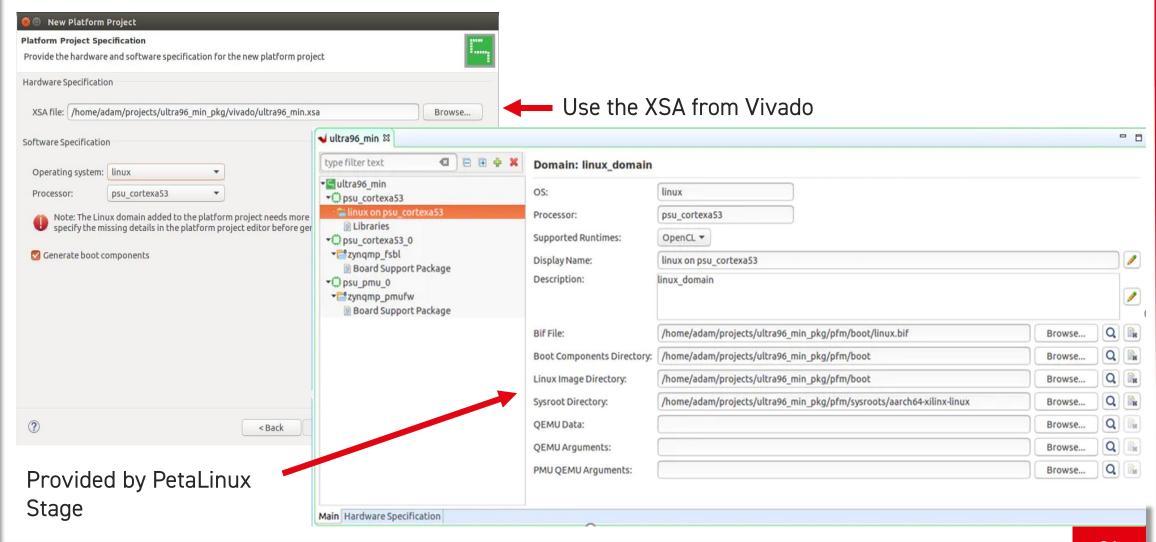




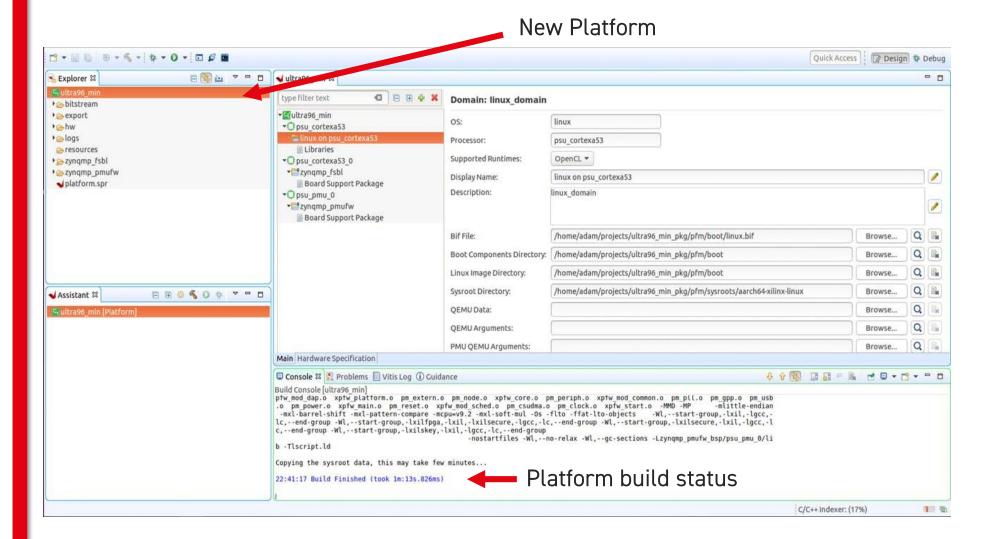
Create a new Platform



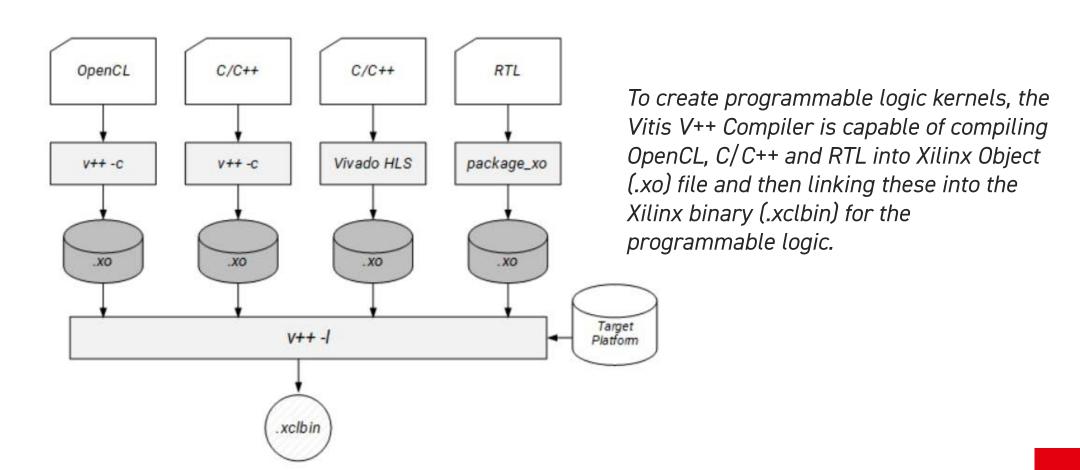




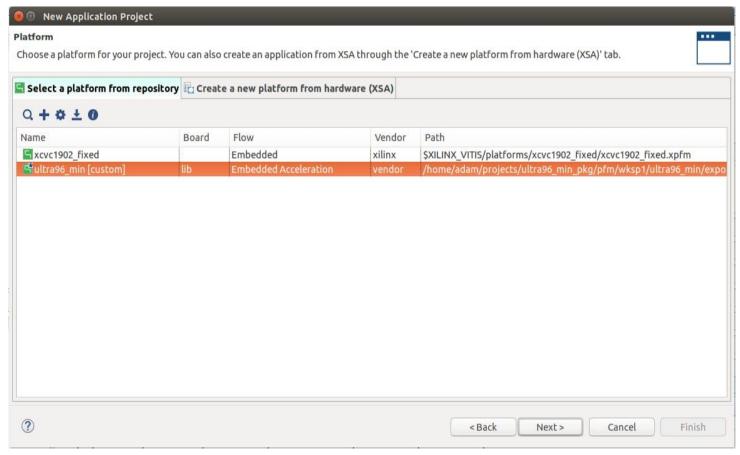






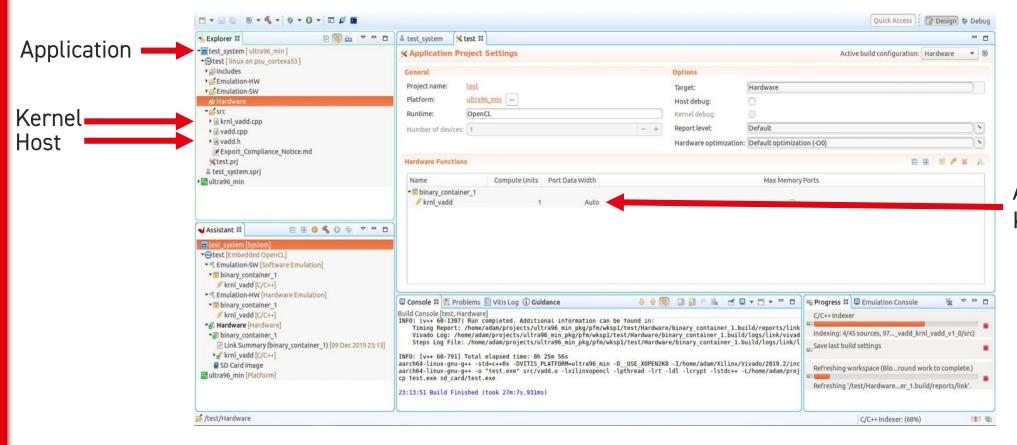






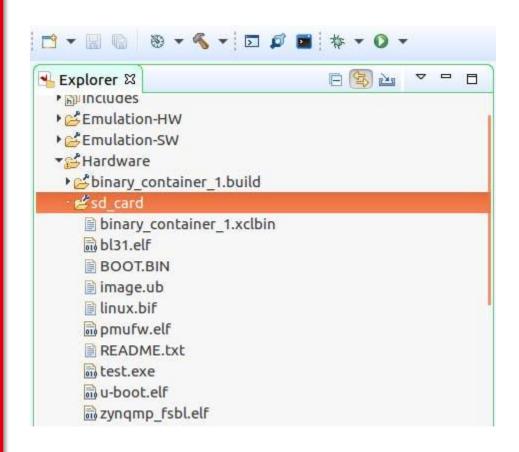
Newly Created Platform





Accelerated Kernel





```
export XILINX_XRT=/usr
```

./test.exe binary_container_1.xclbin







Vitis Example

```
cl::Platform::get(&platforms);
                                                                                                            Find Available Platforms
for(size_t i = 0; (i < platforms.size() ) & (found_device == false) ;i++){</pre>
   cl::Platform platform = platforms[i];
   std::string platformName = platform.getInfo<CL_PLATFORM_NAME>();
   if ( platformName == "Xilinx"){
      devices.clear();
      platform.getDevices(CL_DEVICE_TYPE_ACCELERATOR, &devices);
   if (devices.size()){
      device = devices[0];
                                 // Creating Context and Command Queue for selected device
       found device = true;
                                                                                                                               Create Context
                                 cl::Context context(device);
                                                                                                                               Program
                                 cl::CommandQueue q(context, device, CL QUEUE PROFILING ENABLE);
                                 // Load xclbin
                                 std::cout << "Loading: '" << xclbinFilename << "'\n";</pre>
                                 std::ifstream bin_file(xclbinFilename, std::ifstream::binary);
                                 bin file.seekg (0, bin file.end);
                                 unsigned nb = bin file.tellg();
                                 bin_file.seekg (0, bin_file.beg);
                                 char *buf = new char [nb];
                                 bin_file.read(buf, nb);
                                 // Creating Program from Binary File
                                 cl::Program::Binaries bins;
                                 bins.push back({buf,nb});
                                 devices.resize(1);
                                 cl::Program program(context, devices, bins);
```



```
cl::Buffer buffer a(context, CL MEM READ ONLY, size in bytes);

    Allocate Buffers

 cl::Buffer buffer_b(context, CL_MEM_READ_ONLY, size_in_bytes);
 cl::Buffer buffer result(context, CL MEM WRITE ONLY, size in bytes);
int *ptr_a = (int *) q.enqueueMapBuffer (buffer_a , CL_TRUE , CL_MAP_WRITE ,
0, size in bytes);
                                                                             Map buffers to the kernel
int *ptr_b = (int *) q.enqueueMapBuffer (buffer_b , CL_TRUE , CL_MAP_WRITE ,
0, size in bytes);
int *ptr_result = (int *) q.enqueueMapBuffer (buffer_result , CL_TRUE , CL_MA
P READ , 0, size in bytes);
// Data will be migrated to kernel space
                                                                                   Migrate and execute kernel
q.enqueueMigrateMemObjects({buffer a,buffer b},0/* 0 means from host*/);
 // order to view the results. This call will transfer the data from FPGA to
 // source results vector
```



Vitis Example

```
void krnl vadd(const unsigned int *in1, // Read-Only Vector 1
                                                                               Interfaces
const unsigned int *in2, // Read-Only Vector 2
unsigned int *out_r, // Output Result
int size
                        // Size in integer
                                                                                 Interface Declarations
#pragma HLS INTERFACE m_axi port = in1 offset = slave bundle = gmem 
#pragma HLS INTERFACE m axi port = in2 offset = slave bundle = gmem
#pragma HLS INTERFACE m_axi port = out_r offset = slave bundle = gmem
#pragma HLS INTERFACE s axilite port = in1 bundle = control
#pragma HLS INTERFACE s_axilite port = in2 bundle = control
#pragma HLS INTERFACE s axilite port = out r bundle = control
#pragma HLS INTERFACE s axilite port = size bundle = control
#pragma HLS INTERFACE s axilite port = return bundle = control
```



```
(int i = 0; i < size; i += BUFFER_SIZE) {</pre>
#pragma HLS LOOP TRIPCOUNT min=c len max=c len
int chunk size = BUFFER SIZE;
//boundary checks
if ((i + BUFFER SIZE) > size)
    chunk size = size - i;
read1: for (int j = 0; j < chunk size; j++) {</pre>
    #pragma HLS LOOP_TRIPCOUNT min=c size max=c size
    #pragma HLS PIPELINE II=1
    v1 buffer[j] = in1[i + j];
//Burst reading B and calculating C and Burst writing
// to Global memory
vadd writeC: for (int j = 0; j < chunk size; j++) {</pre>
    #pragma HLS LOOP TRIPCOUNT min=c size max=c size
    #pragma HLS PIPELINE II=1
    //perform vector addition
    out_r[i+j] = v1_buffer[j] + in2[i+j];
```

Pragma Trip Count = enables HLS tool to understand the number of loop iterations

Pragma Pipeline Initiation Interval = the target number of clocks between being able to process new inputs



Questions?

