PYNQ for HPC

Graham Schelle
Fellow, Adaptive and Embedded Computing Group (AECG)

Agenda

Why Productivity Software for H^2RC?

…a Machine Learning example (DPU-PYNQ)

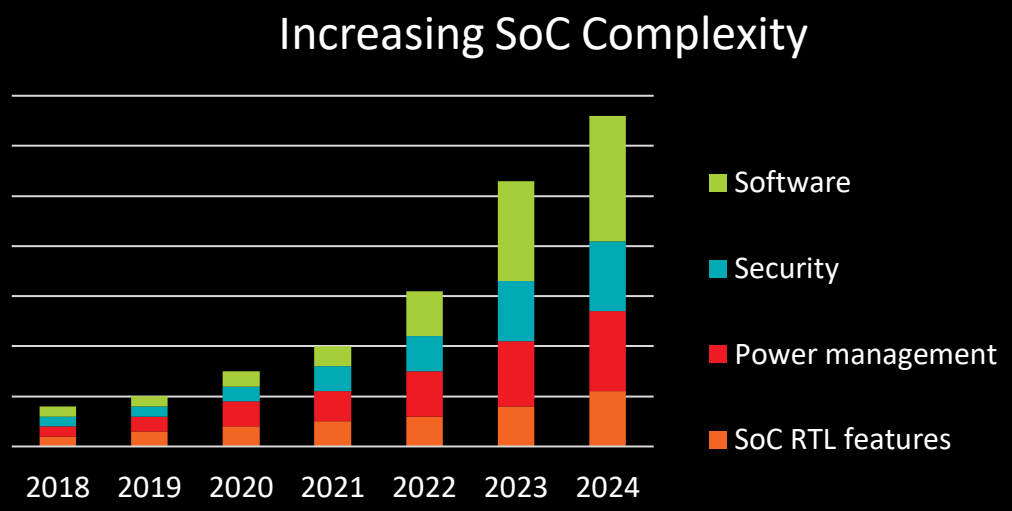
…a Cloud example (PYXRT)

…a Quantum example (QICK)

And what about Productivity Software for Hardware?

…PYNQ-Metadata

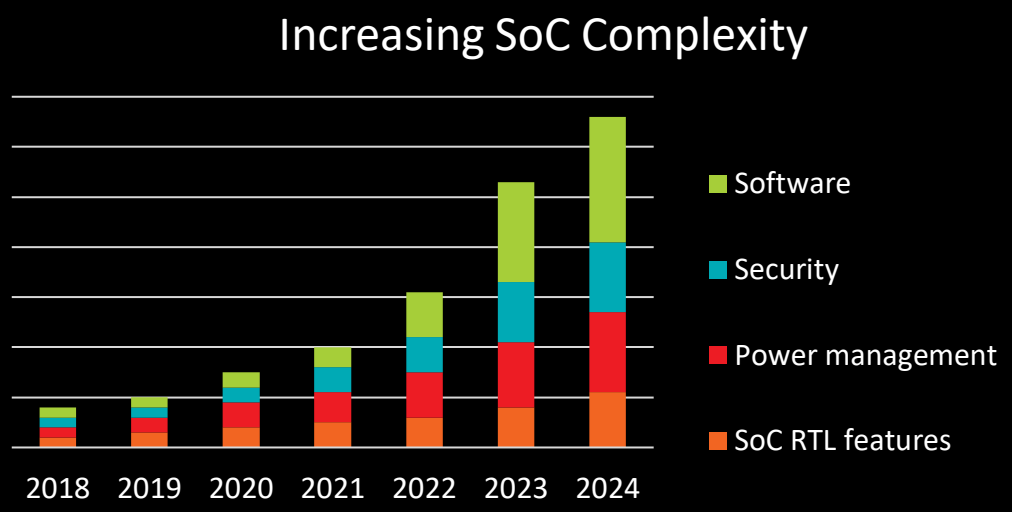
Why Productivity Software for H^2RC?



H^2 : Heterogeneous and High Performance

How can productivity software help here?

Why Productivity Software for H^2RC?



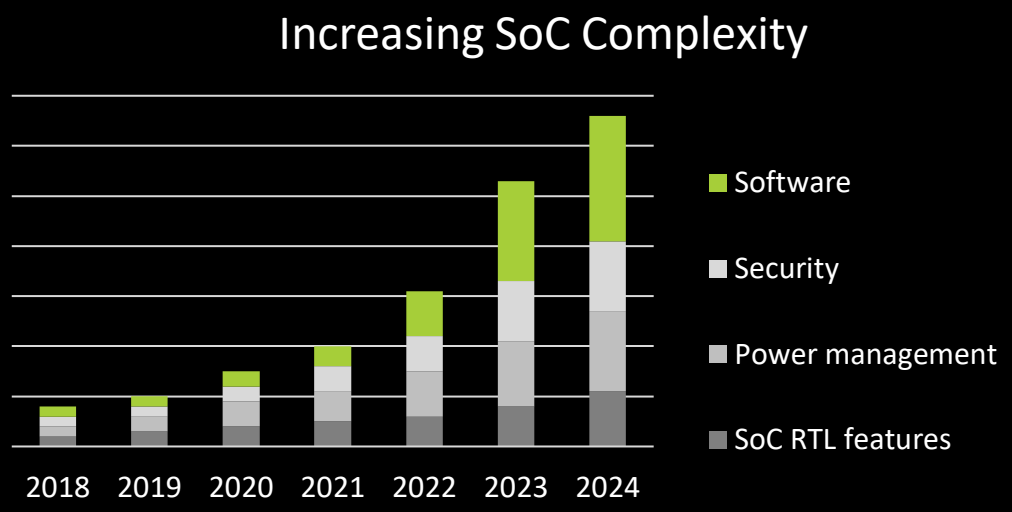
H^2 : Heterogeneous and High Performance

Hardware complexity often exposed to software

…to enable user programmability

→ Software productivity challenges (and opportunities!)

Why Productivity Software for H²RC?



What is the “it” in this talk?

Glue It All Together With Python

Python is an advanced scripting language that is being used successfully to glue together large software components. It spans multiple platforms, middleware products, and application domains. Python is an object-oriented language with high-level data structures, dynamic typing, and dynamic binding. Python has been around since 1991, and has a very active user community. For more information, see the Python website http://www.python.org.

Guido van Rossum, 1998

AMD

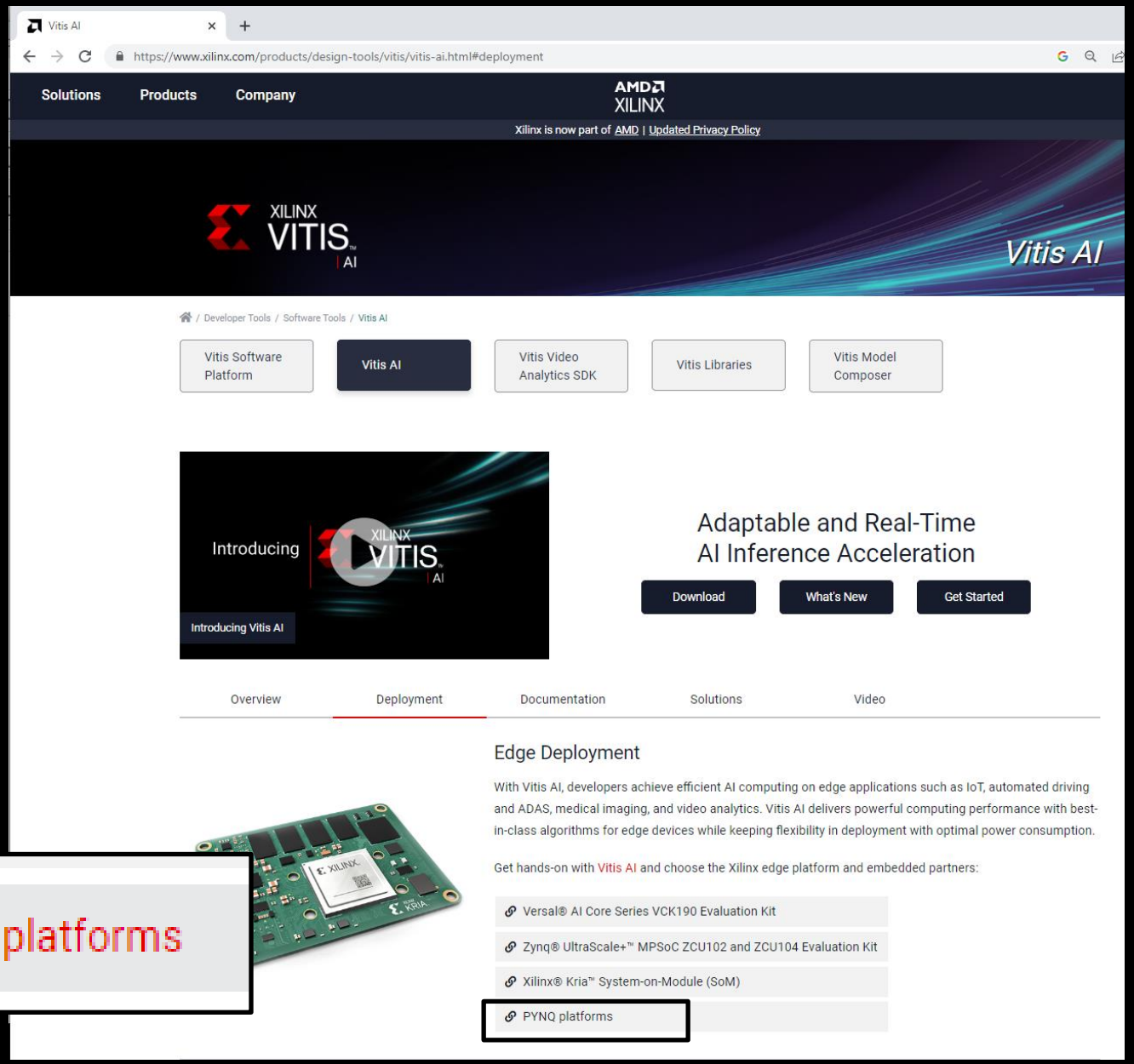
Example – DPU-PYNQ

The DPU is AMD's Adaptive Computing Deep Learning Accelerator

Leverages a host-machine quantization framework for training and then a target runtime framework for execution

Its an IP block found within Vivado, instantiated within VitisAI, user software written in Vitis

Deployed on Zynq UltraScale, Versal & Alveo



Jupyter Notebook
Running DPU-PYNQ

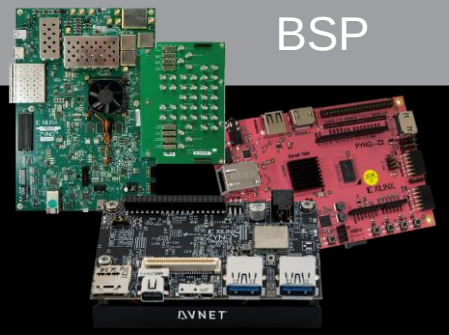
apt/pip | overlays | pynq

PYNQ Virtual Env (venv)

PYNQ Linux Rootfs

Kernel / Drivers

BSP



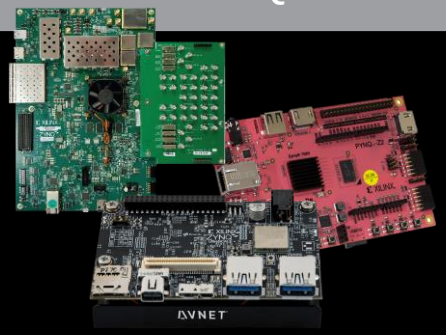
Jupyter Notebook
Running DPU-PYNQ

`pip install dpu-pynq`

pynq

PYNQ venv

PYNQ Linux

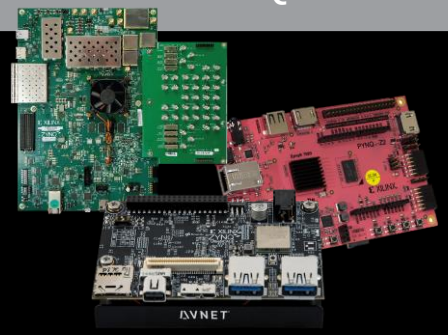


`pip install dpu-pynq`

pynq

PYNQ venv

PYNQ Linux



Jupyter Notebook
Running DPU-PYNQ

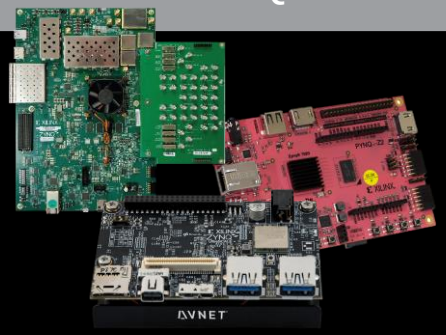
```
from pynq_dpu import DpuOverlay
overlay = DpuOverlay("dpu.bit")

overlay.load_model("dpu_resnet50.xmodel")
```

`pip install dpu-pynq`

pynq

PYNQ venv

PYNQ Linux



Xilinx / **DPU-PYNQ** Public

Xilinx / **PYNQ** Public

Avnet / **Ultra96-PYNQ** Public

MicroPhase / **antsdr-pynq** Public

aldec / **TySOM-PYNQ** Public

PYNQ v3.0 is now available. We've added pynq-metadata with support for hardware design introspection, more supported boards like the #zcu208 and we continue to actively maintain PYNQ-Linux with upgraded software packages.

PYNQ™

PYNQ v3.0 (Belfast) Release

discuss.pynq.io • 1 min read

Jeff Fifield and 202 others — 11 comments • 7 reposts

Like | Comment | Repost | Send

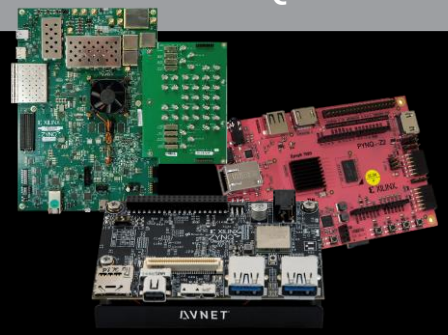
11,213 impressions — View analytics

`pip install dpu-pynq`

pynq

PYNQ venv

PYNQ Linux



DPU-PYNQ with Vitis AI 2.5 support now available... but more importantly now available on most platforms we could find.

PYNQ™

DPU On 20+ boards

DPU-PYNQ v2.5.0 release -- now on 20+ boards

discuss.pynq.io • 1 min read

Cathal McCabe and 122 others — 2 comments • 7 reposts

Like | Comment | Repost | Send

5,324 impressions — View analytics

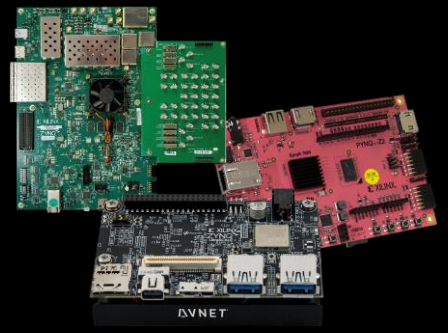
AMD

Edge to Cloud Productivity

pynq

PYNQ venv

PYNQ Linux



https://github.com/Xilinx/PYNQ/blob/master/sdbuild/packages/python_packages_jammy/requirements.txt

Search or jump to... / Pulls Issues Marketplace Explore

Xilinx / PYNQ Public

Edit Pins | Unwatch 128 | Fork 732 | Starred 1.5k

Code | Issues 19 | Pull requests 7 | Discussions | Actions | Projects

master | PYNQ / sdbuild / packages / python_packages_jammy / requirements.txt | Go to file

skalade packages updated to be compatible with ubuntu 22.04 and ... Latest commit f8381e6 on Sep 28 History

1 contributor

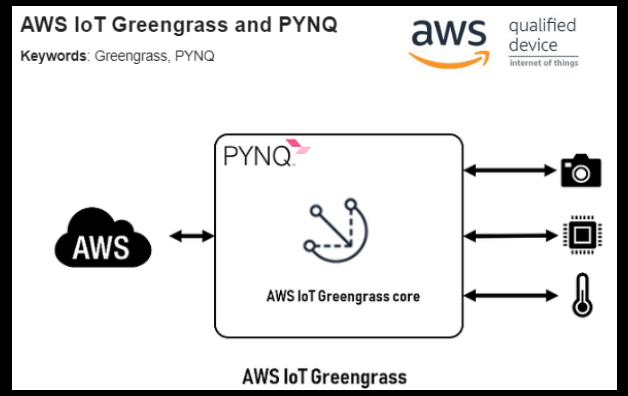
Executable File | 123 lines (123 sloc) | 2.35 KB | Raw | Blame

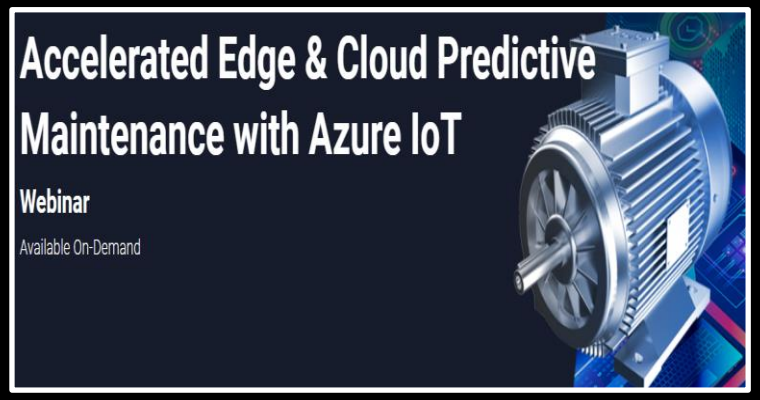
```
1  alabaster==0.7.12
2  anyio==3.6.1
3  argon2-cffi==21.3.0
4  argon2-cffi-bindings==21.2.0
5  asttokens==2.0.5
```

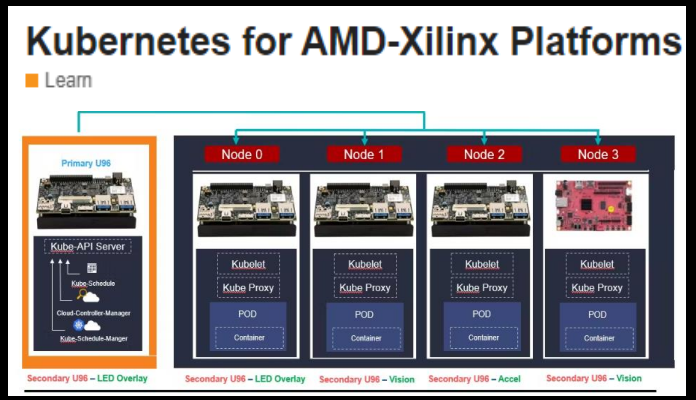
pip install -r requirements.txt
pip install pynq

AMD

Edge to Cloud Productivity



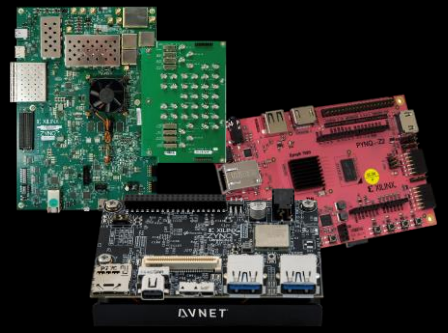




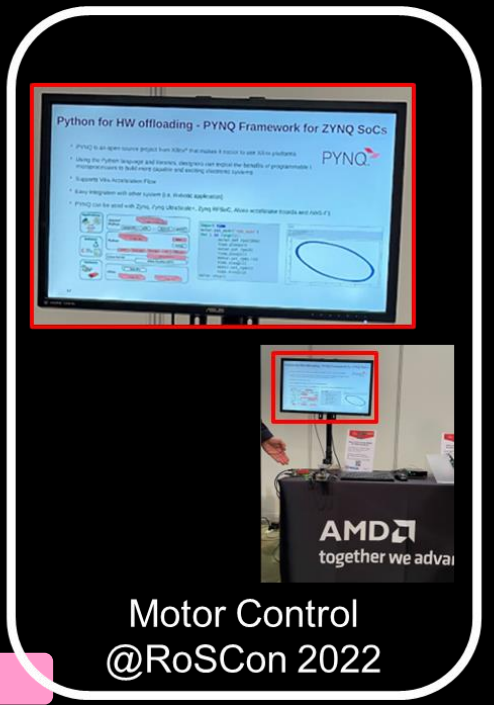
pynq

PYNQ venv

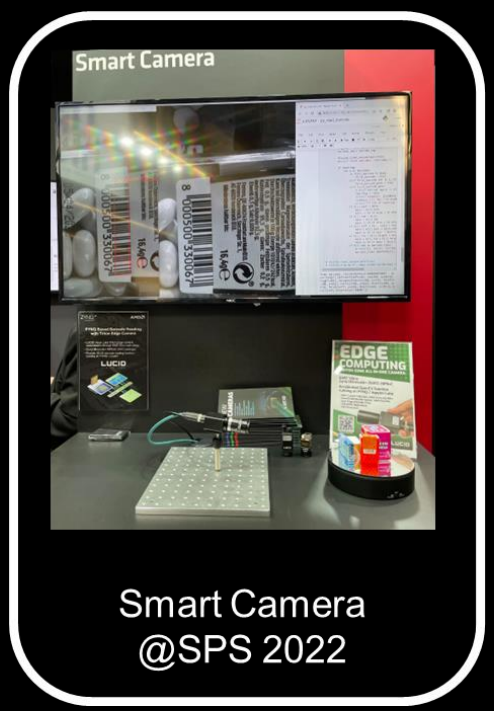
PYNQ Linux



Edge to Cloud Productivity



Motor Control
@RoSCon 2022



Smart Camera
@SPS 2022

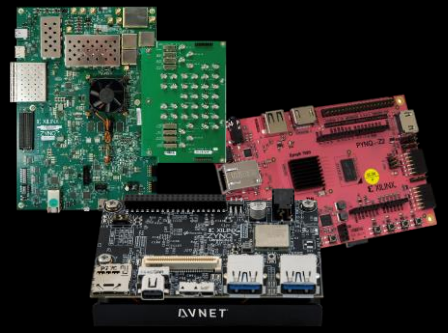
pynq

PYNQ venv

Kria-PYNQ/install.sh

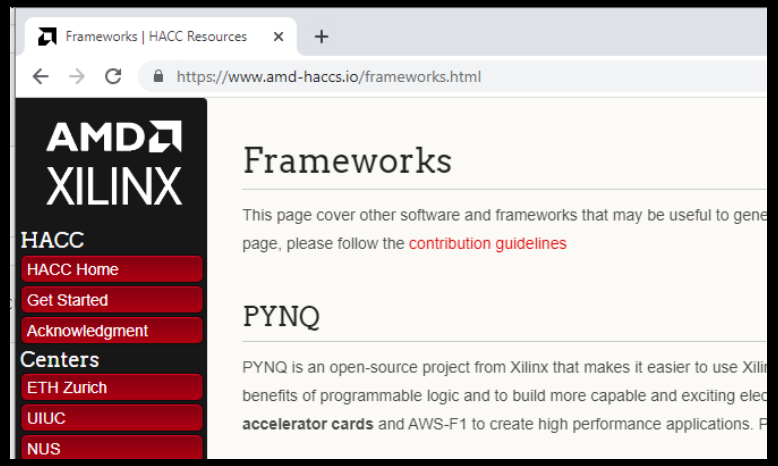
PYNQ-Linux

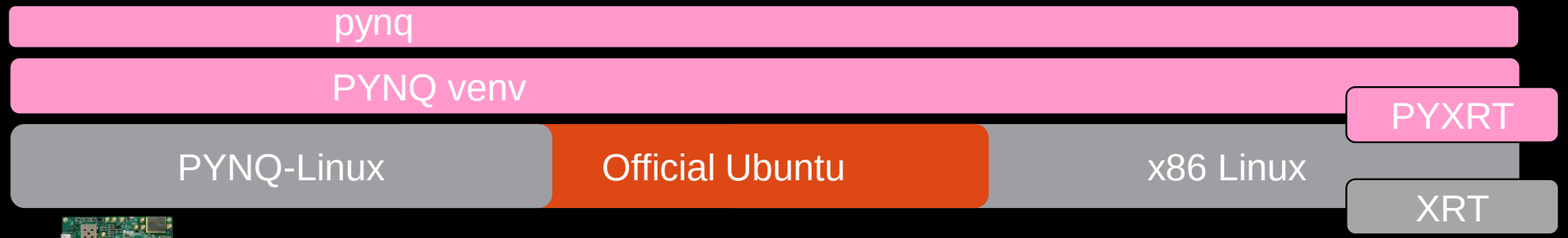
Official Ubuntu

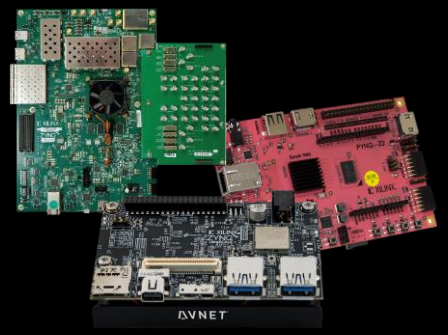


AMD

Edge to Cloud Productivity







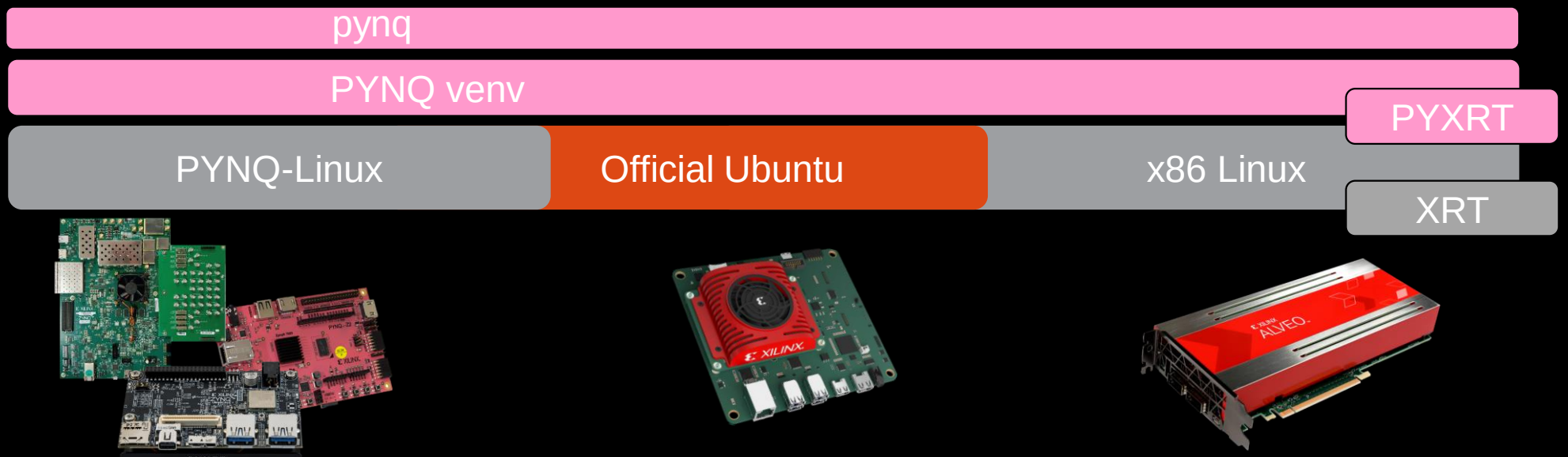
AMD

Edge to Cloud Productivity

```
# which platforms to build the [neural] networks for
zynq_platforms = ["ZCU102", "ZCU104"]
alveo_platforms = ["U250"]
platforms_to_build = zynq_platforms + alveo_platforms
```

Xilinx / finn-examples Public

```
ol = Overlay("an_example.bit")
ol = Overlay("an_example.xclbin")
ol = Overlay("an_example.xsa")

ol = Overlay("an_example.bit", dtbo="a_devicetreeoverlay.dtbo")
```



AMD

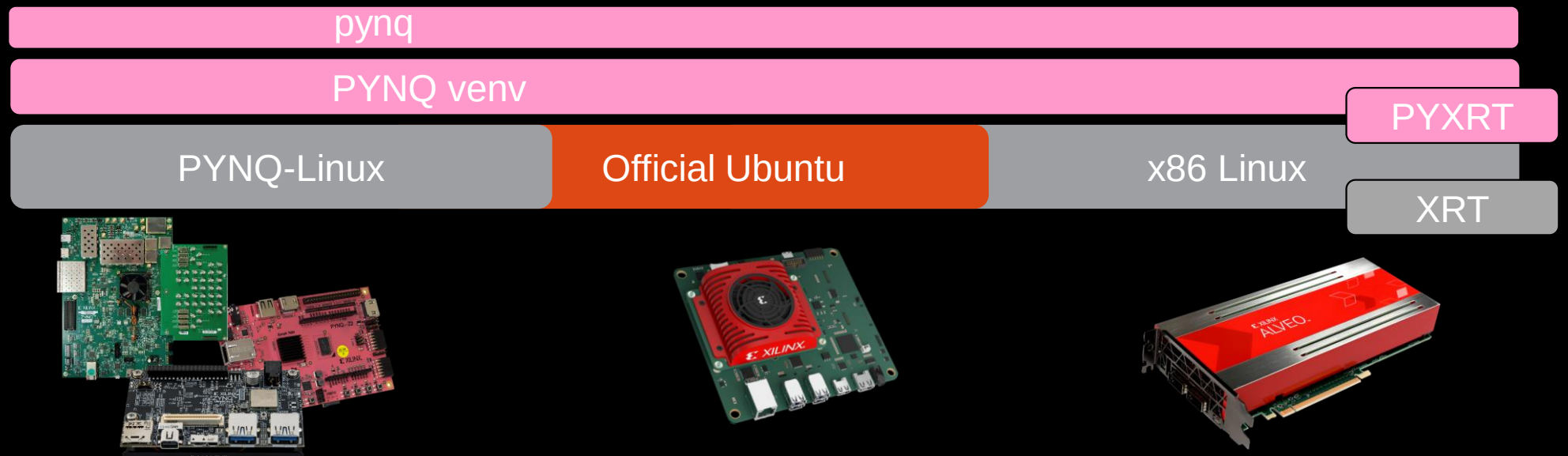
Edge to Cloud Productivity

PyLog: An Algorithm-Centric Python-Based FPGA Programming and Synthesis Flow

Sitao Huang, Kun Wu, Hyunmin Jeong, Chengyue Wang, Deming Chen, *Fellow, IEEE*, and Wen-Mei Hwu, *Fellow, IEEE*

PyLog runtime uses PYNQ [20] and XRT, which are runtime libraries from Xilinx. The lower half of Fig. 2 shows an example of PCIe-based FPGA platform. Note that PyLog can support both PCIe-based high-performance FPGAs and low-power SoCs and MPSoCs

Unified edge to cloud API
for adaptive computing



AMD

Edge to Cloud Productivity

```
buffer = pynq.allocate(shape=(height, width, 3), dtype=np.uint8)
```

numpy

cache flush/invalidate (Zynq)

Migrate to/from device (Opencl)

pynq

PYNQ venv

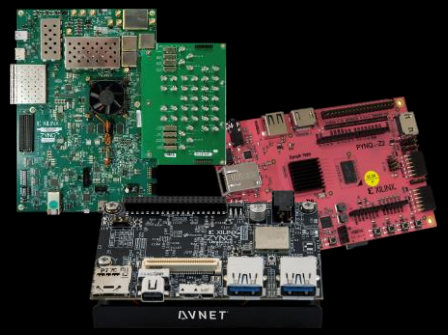
PYXRT

PYNQ-Linux

Official Ubuntu

x86 Linux

XRT



AMD

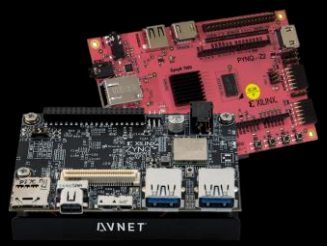
The New Cloud - Quantum

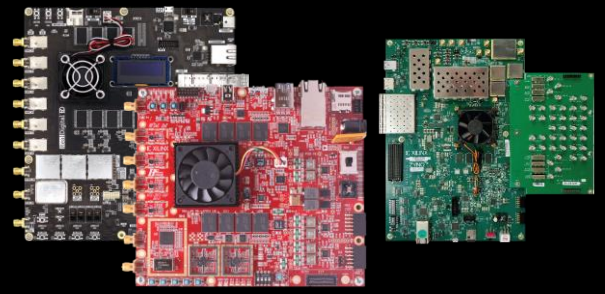
RFSoC-PYNQ

pynq

PYNQ venv

PYNQ-Linux





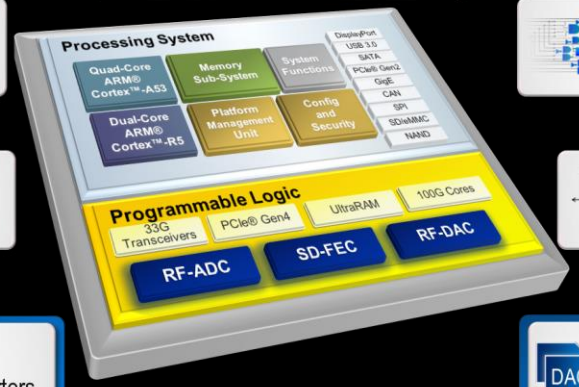
ARM Cortex
Processing System
- Quad-Core A53 (64-bit)
- Dual-Core R5 (32-bit)

Hardened Engines
- PCIe Gen3 & Gen4
- 100G Cores

Programmable Logic
- 16nm FinFET
- UltraScale+ FPGA Fabric

33G Transceivers
- 33Gb/s
- 28G Backplane Capable

DSP-Intensive
- 4,272 DSP slices
- 7,612 GMACs



ADC
Integrated Direct-RF
Analog-to-Digital Converters

DAC
Integrated Direct-RF
Digital-to-Analog Converters

Soft Decision Forward
Error Correction
LDPC & Turbo Support

AMD

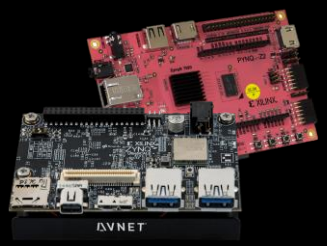
The New Cloud - Quantum

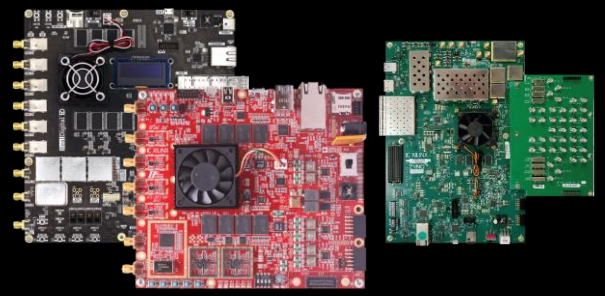
RFSoC-PYNQ

pynq

PYNQ venv

PYNQ-Linux





strath-sdr / **rfsoc_sam** Public

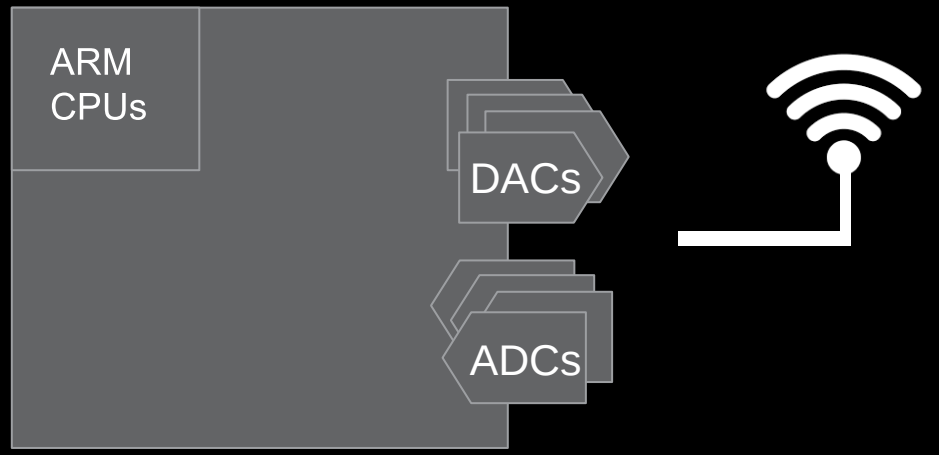
Xilinx / **DSP-PYNQ** Public

Xilinx / **SDFEC-PYNQ** Public

AMD

The New Cloud - Quantum

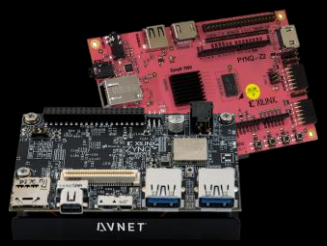
RFSoC-PYNQ

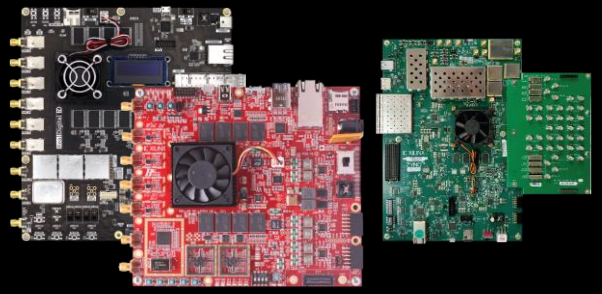


pynq

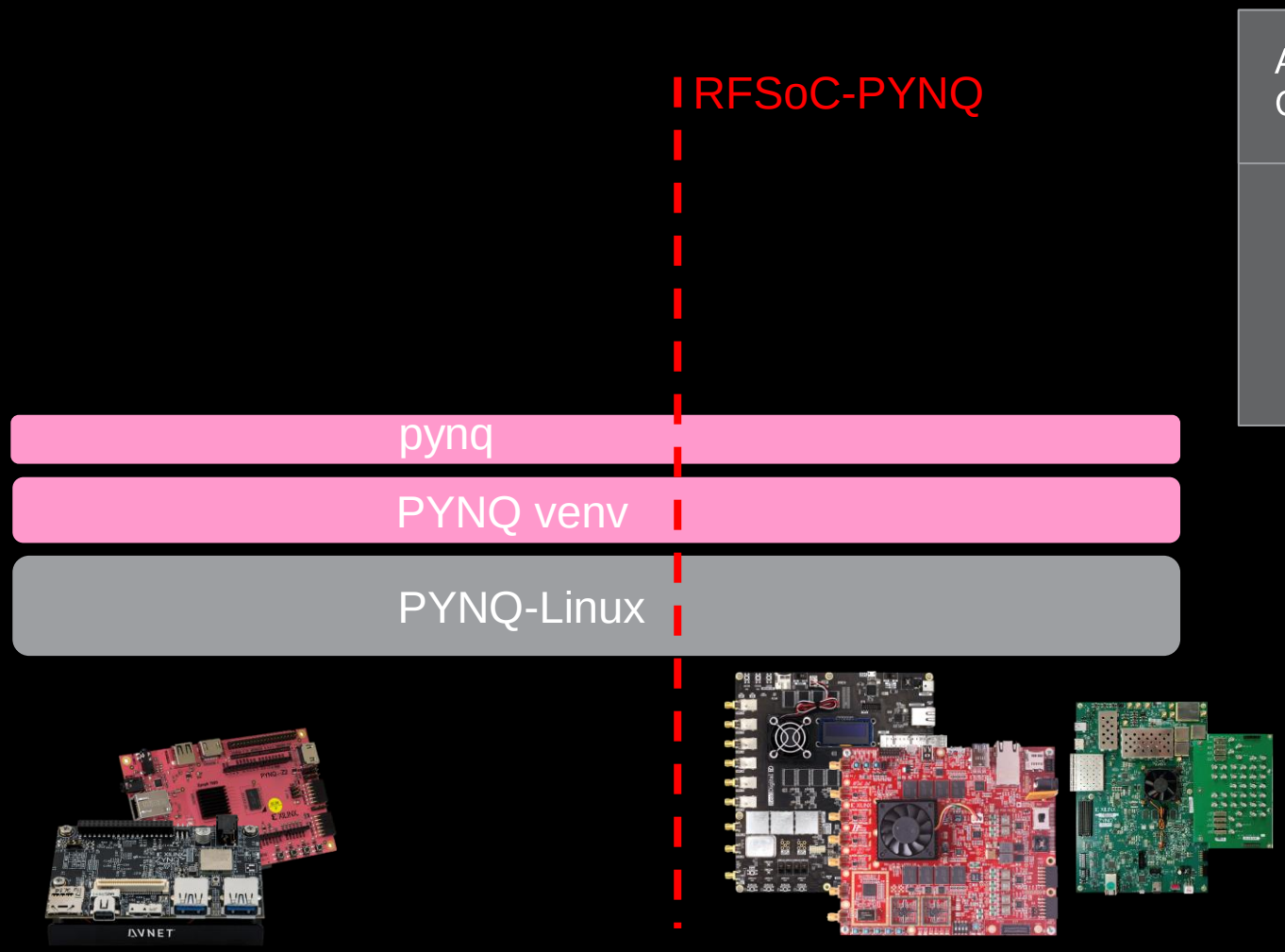
PYNQ venv

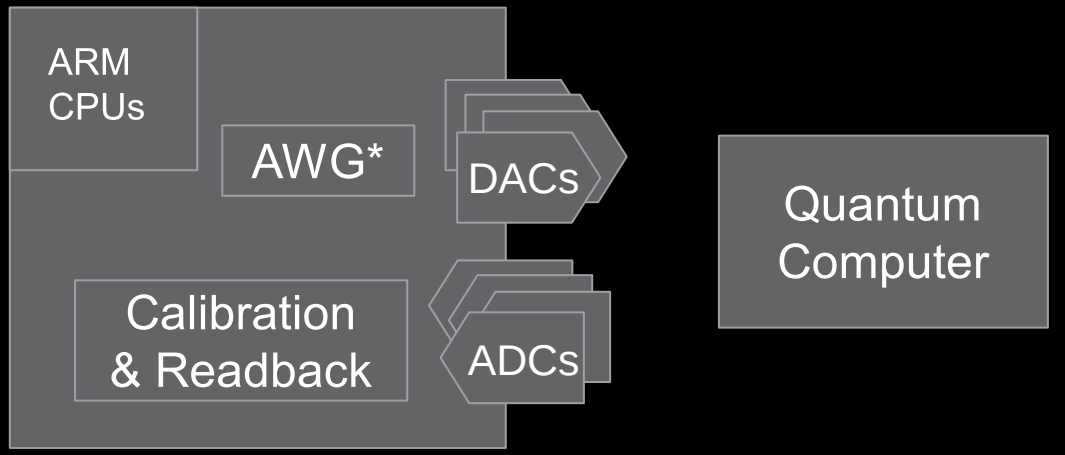
PYNQ-Linux





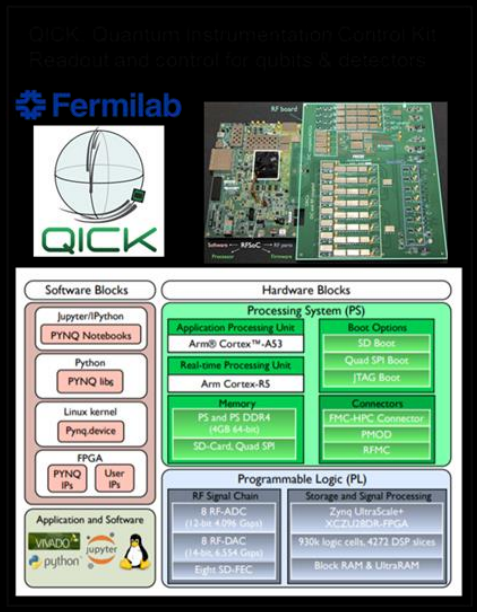
The New Cloud - Quantum

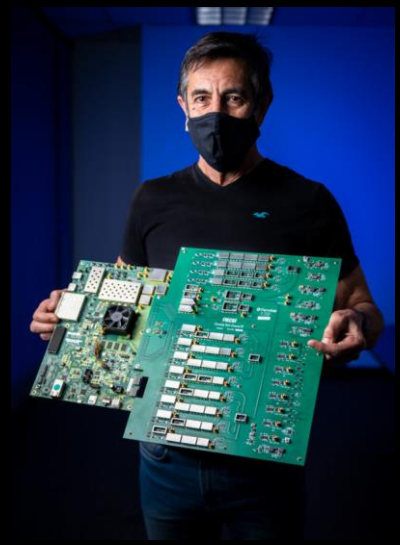




*AWG – Arbitrary Waveform Generator

The New Cloud - Quantum



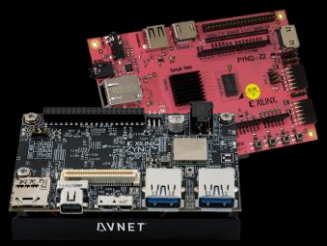


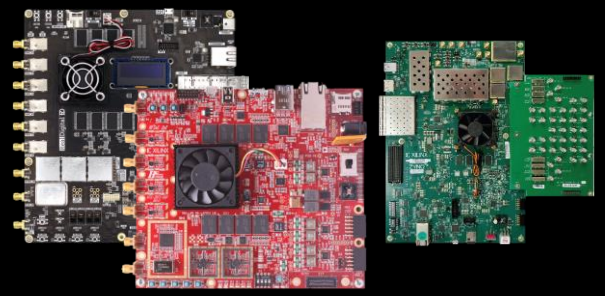
RFSoC-PYNQ

pynq

PYNQ venv

PYNQ-Linux





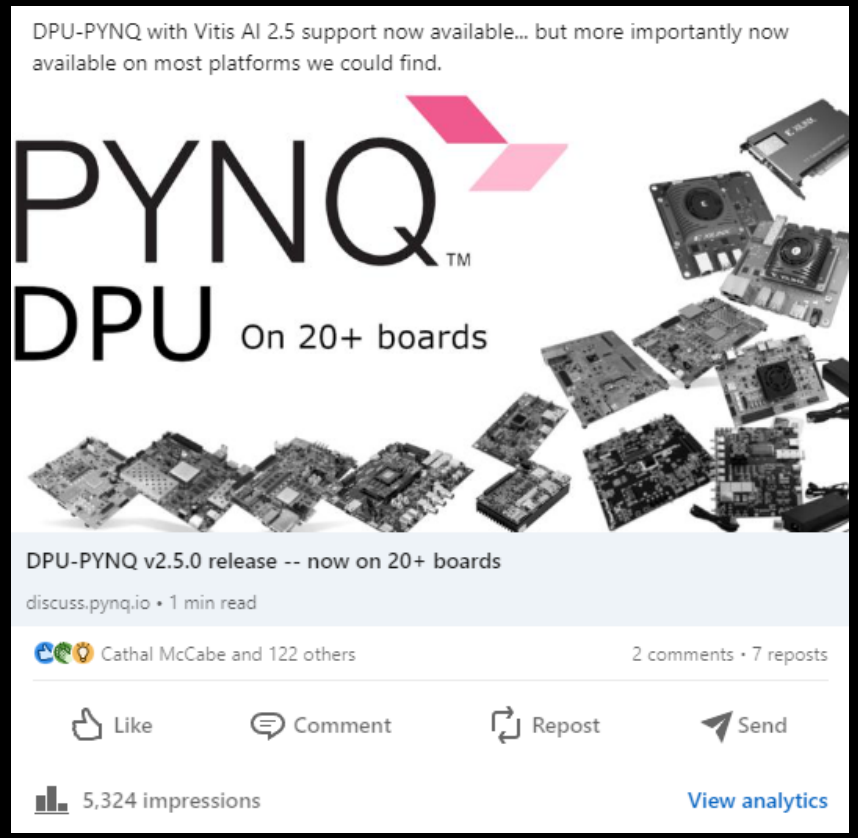
Gustavo Cancelo led a team of Fermilab engineers to create a new compact electronics board: It has the **capabilities of an entire rack of equipment** that is compatible with many designs of superconducting qubits at a **fraction of the cost**.

HPCwire.com

openquantumhardware / qick Public

AMD

Coming Soon … PYNQ-Metadata for Hardware Reuse




```python
from pynqutils import Boards, VitisAI
from pynqmetadata import Metadata

zu_boards = Boards(families=['ZynqUltraScale'])

for b in zu_boards:
    vai_overlay = VitisAI( md=Metadata('zcu104_dpu.hwh'), board=b )
    vai_overlay.generate()
```

1. Boards and existing designs as Python objects
2. Declarative manipulations to change the design
3. Simple backend to reconstruct the design

AMD

Summary

Glue It All Together With Python

Glue Heterogeneous Development and Integration Together with Productivity Frameworks

E.g. PYNQ

Nothing shown gets in the way of “High Performance” …

Everything shown is to simplify “Reconfigurable Computing” …

Everything shown is open source, community supported

Please contact me to collaborate on how we can make Reconfigurable Computing more accessible to the HPC community

graham.schelle@amd.com

Thank you

Please contact me to collaborate on how we can make Reconfigurable Computing more accessible to the HPC community

graham.schelle@amd.com

AMD


AMD