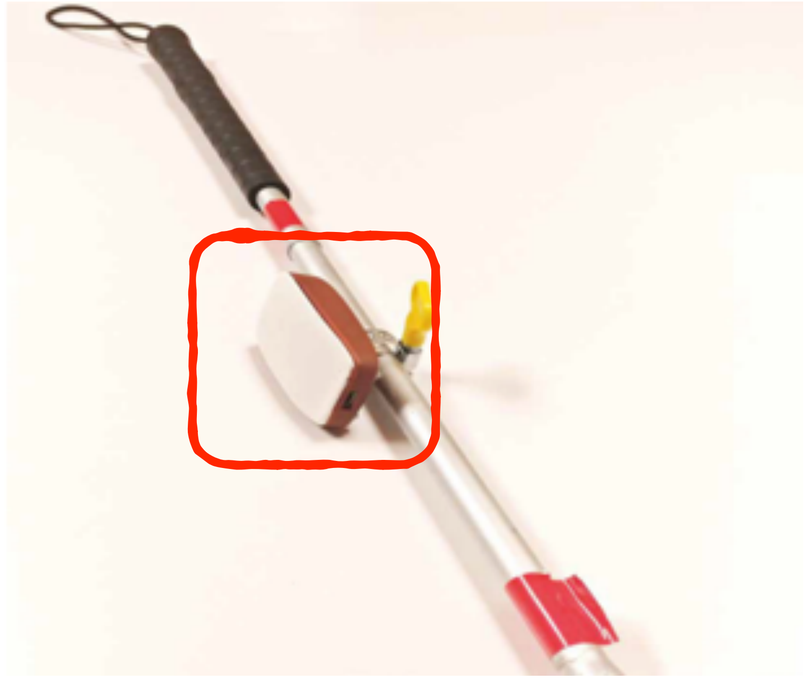
µTVM
Deep Learning on Bare-Metal Devices

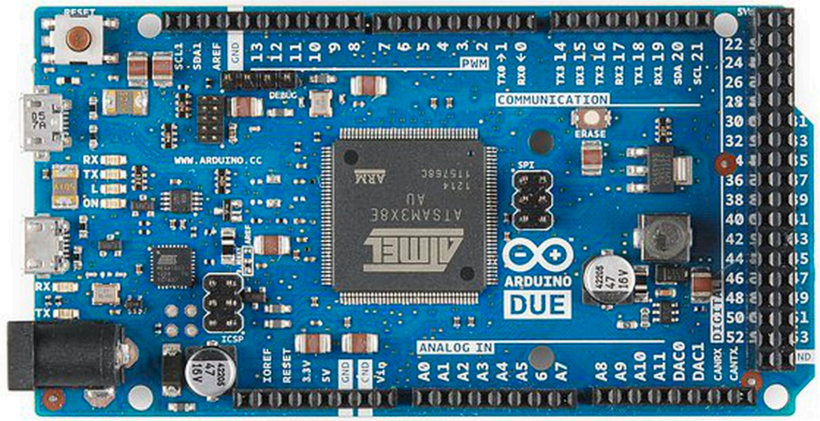
Pratyush Patel

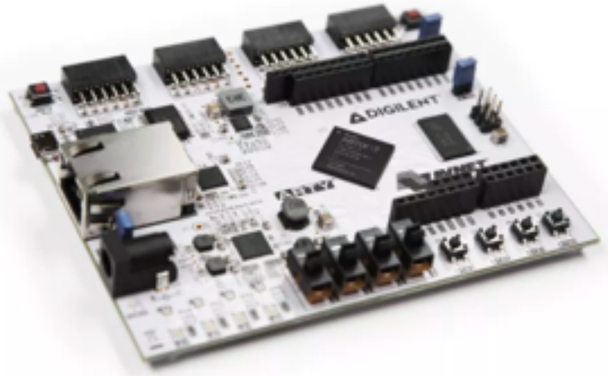
sampl

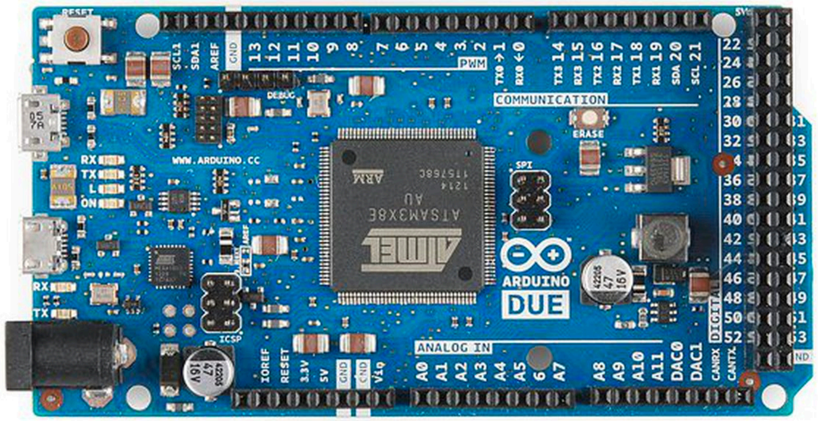
PAUL G. ALLEN SCHOOL
OF COMPUTER SCIENCE & ENGINEERING

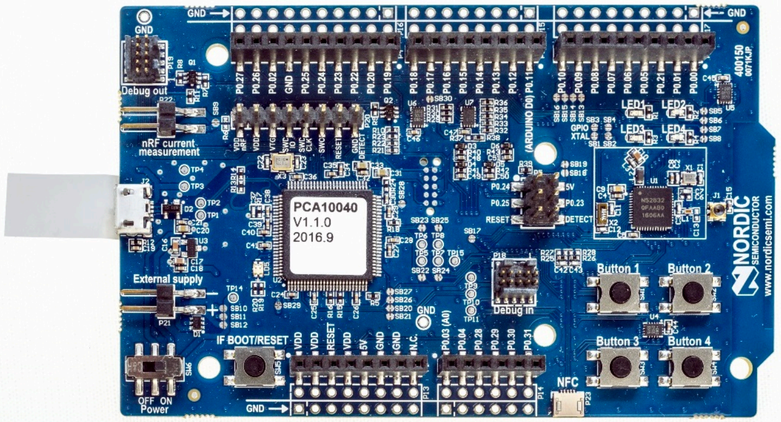
tvm.ai

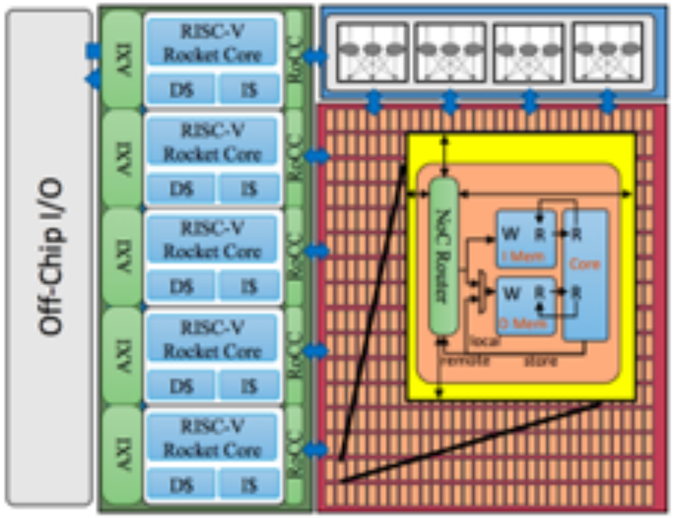



RESET
PWM
COMMUNICATION
ERASE
SPI
WWW.ARDUINO.CC
ATSAM3X8E
ARDUINO
DUE
ANALOG IN
DIGITAL

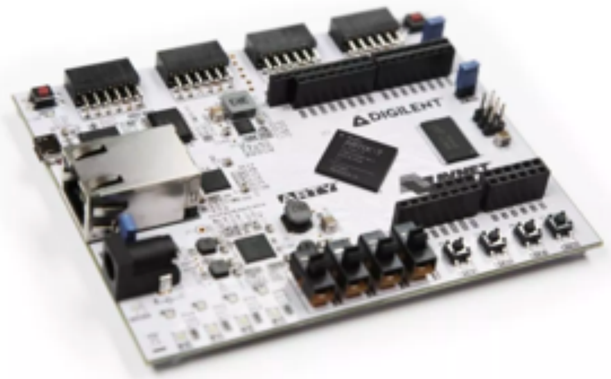

DIGILENT

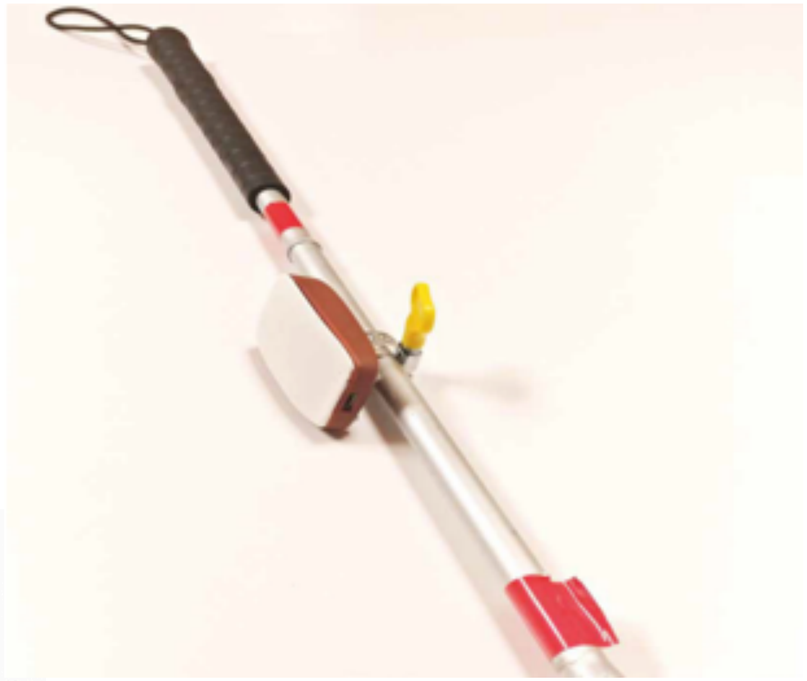

ARDUINO
DUE
WWW.ARDUINO.CC
COMMUNICATION
PWM
ANALOG IN
ATMEL

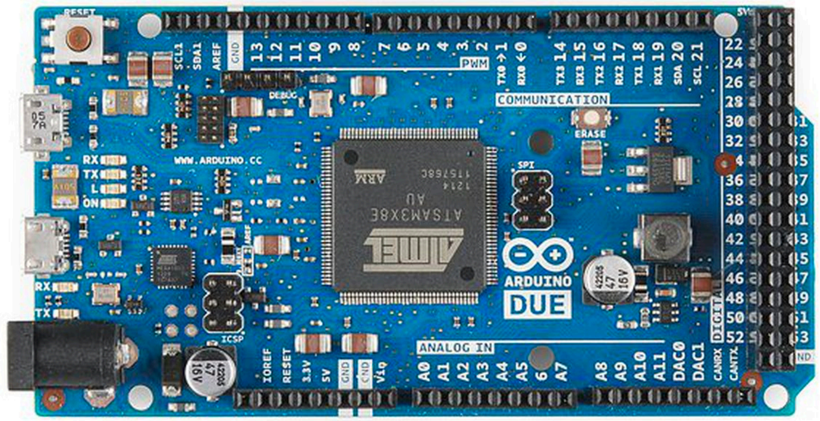

PCA10040
V1.1.0
2016.9
NORDIC SEMICONDUCTOR
www.nordicsemi.com
Debug out
nRF current measurement
External supply
IF BOOT/RESET
OFF ON Power
Debug in
NFC
Button 1
Button 2
Button 3
Button 4
LED1
LED2
LED3
LED4

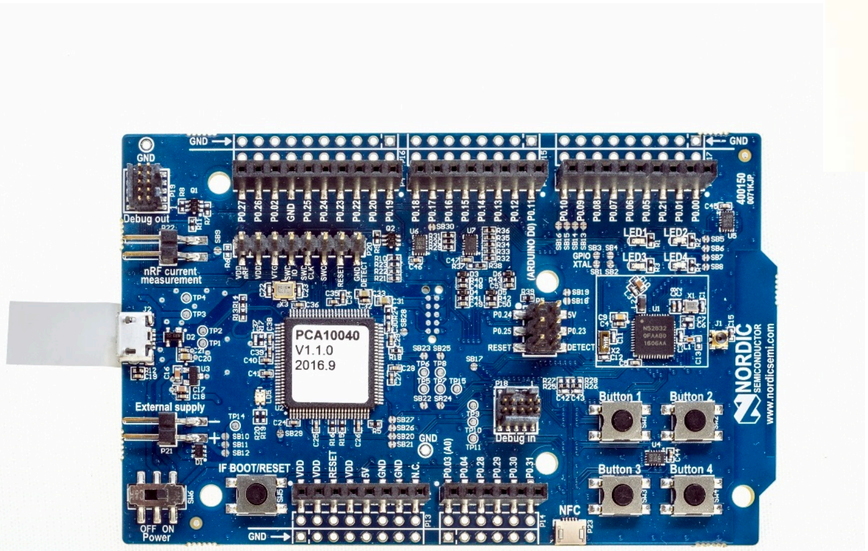

Off-Chip I/O
AXI
RISC-V Rocket Core
D$
I$
NoC Router
I Mem
D Mem
Core
local
remote
store

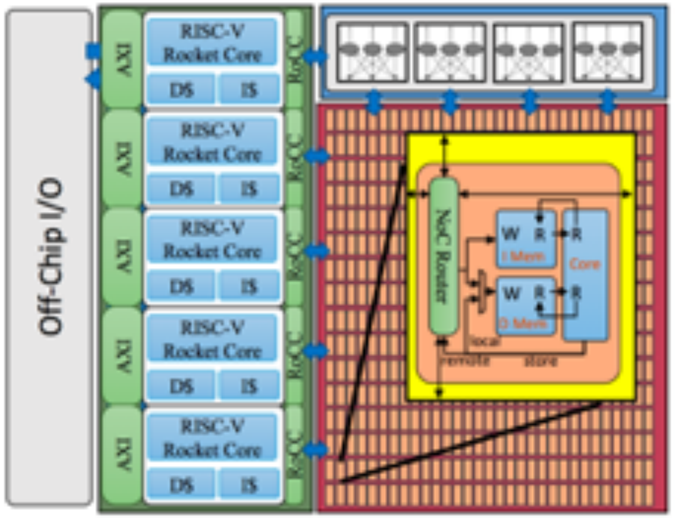
No OS stack



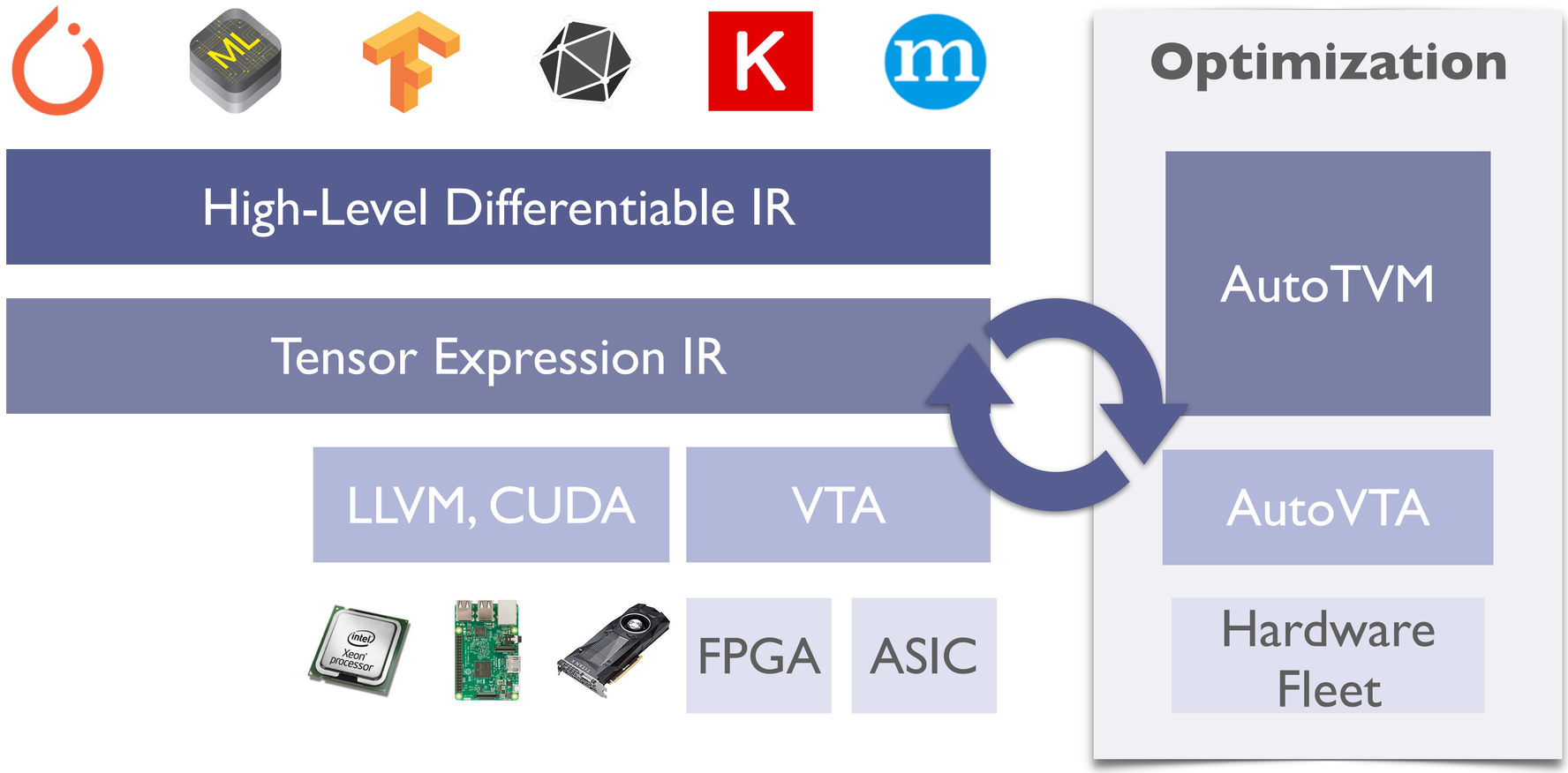



COMMUNICATION
ANALOG IN
ARDUINO
DUE

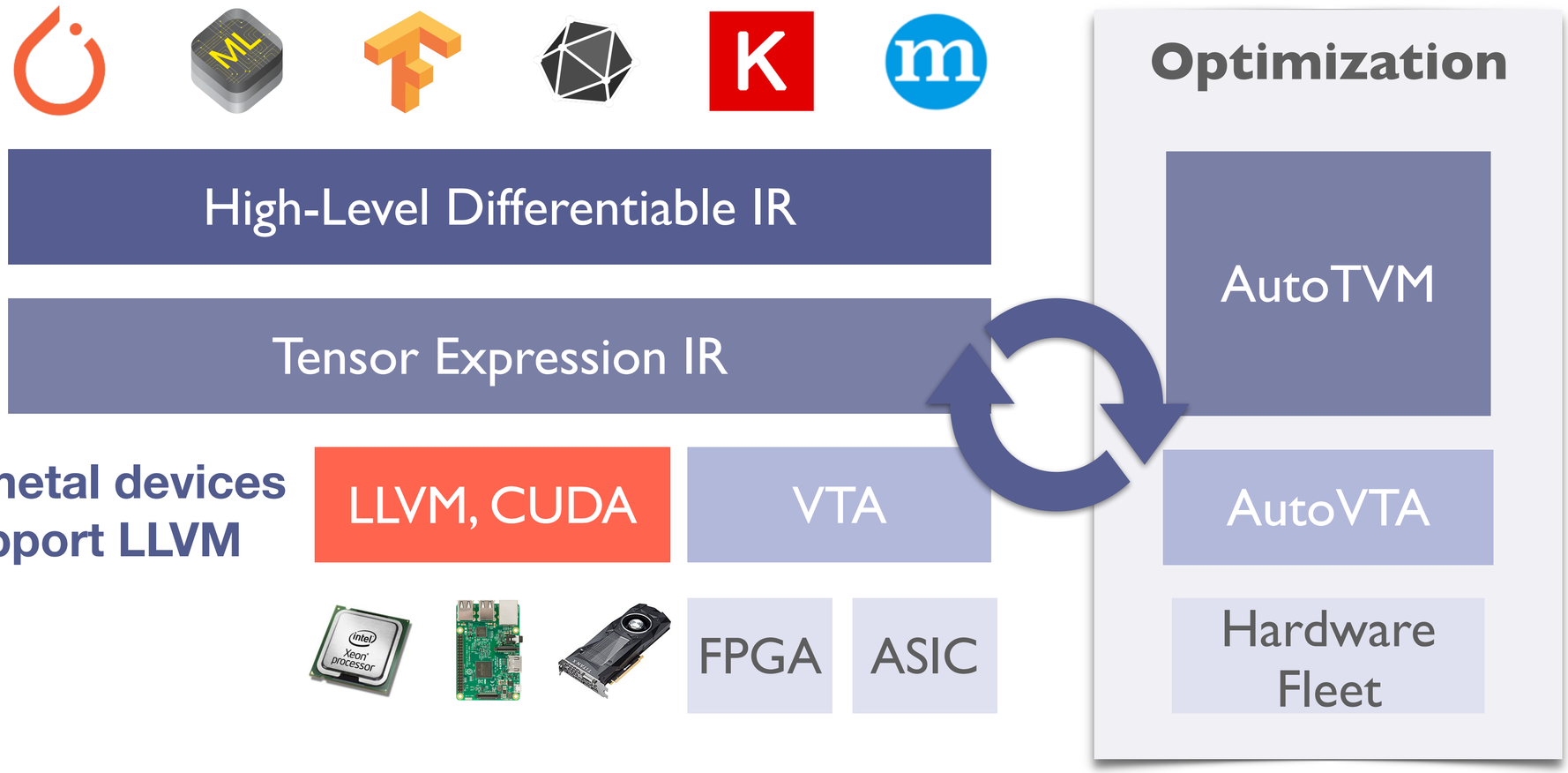

PCA10040
V1.1.0
2016.9
Button 1
Button 2
Button 3
Button 4
NORDIC
SEMICONDUCTOR


Off-Chip I/O
AXI
RISC-V
Rocket Core
DS
IS
RoCC
NoC Router
Core

Extend TVM to support bare-metal devices

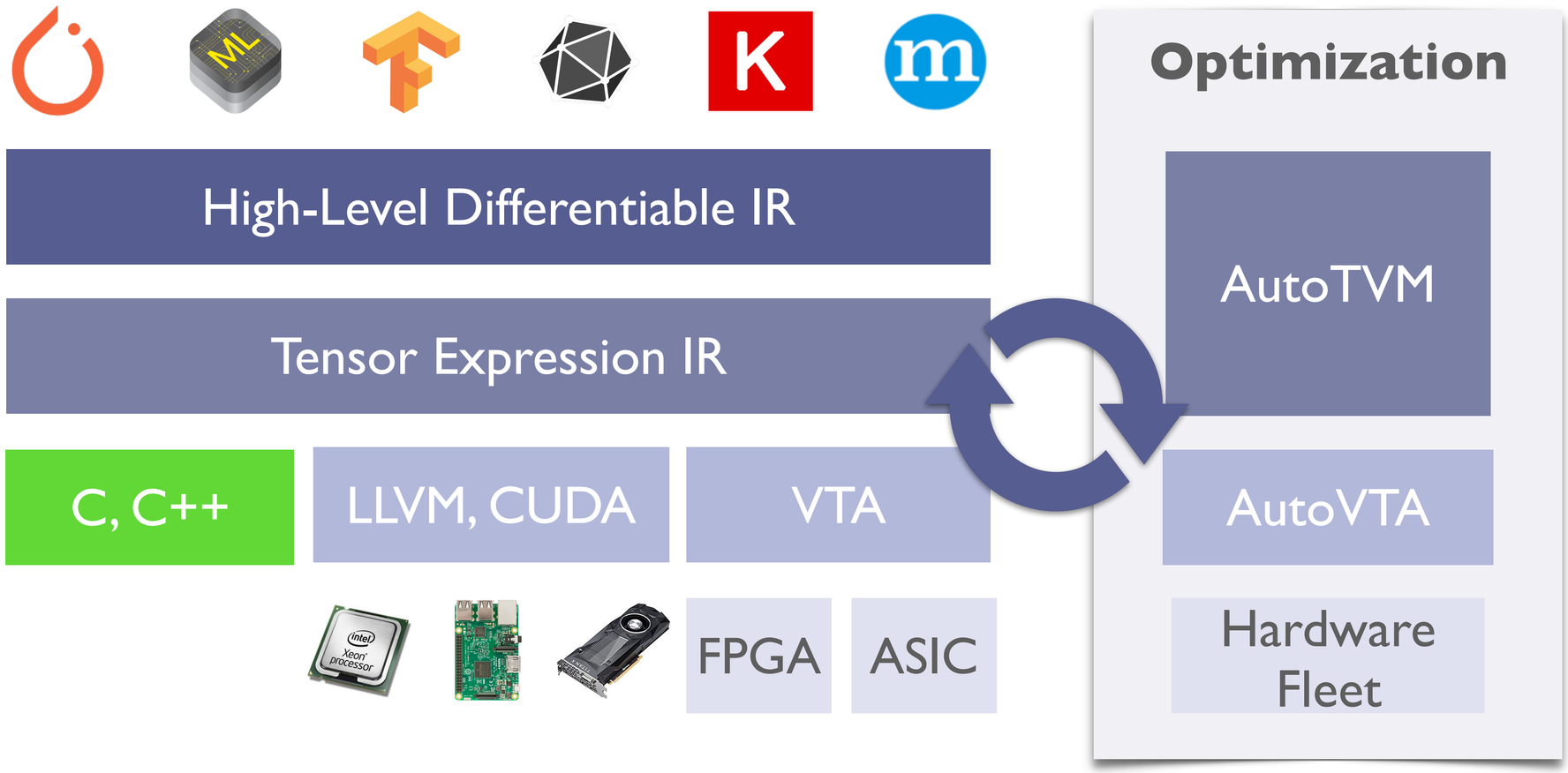


Extend TVM to support bare-metal devices

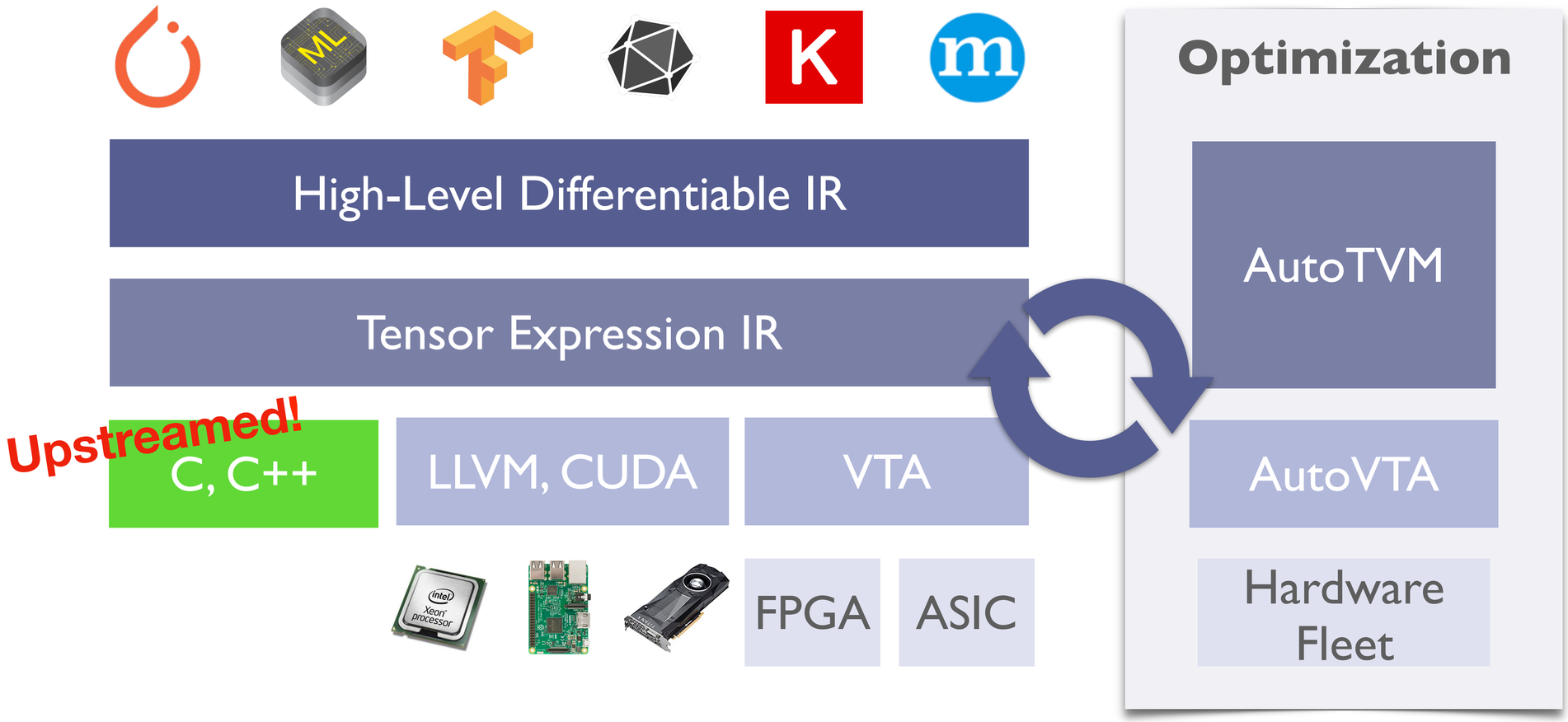


Most bare-metal devices do not support LLVM

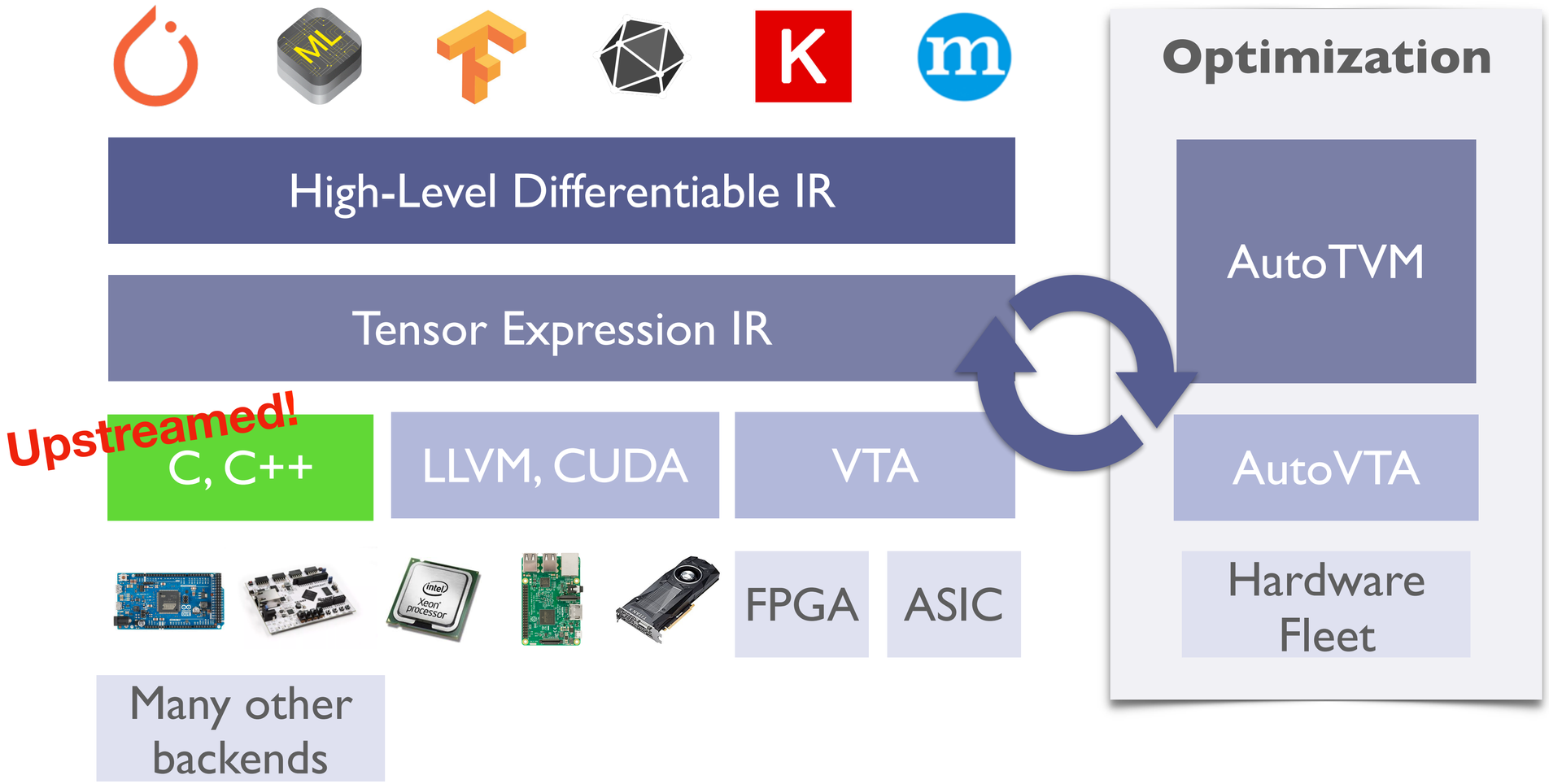
Extend TVM to support bare-metal devices



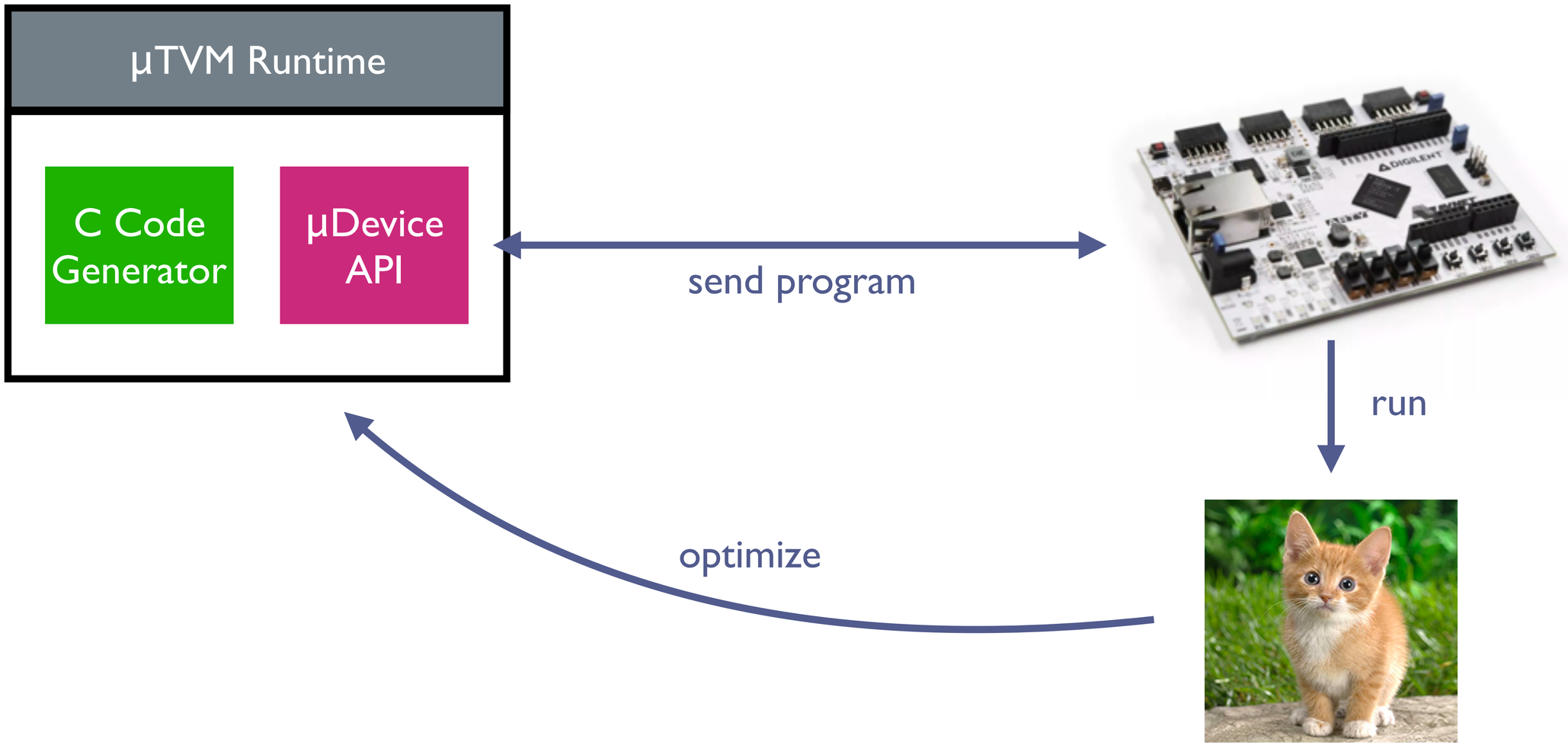
Extend TVM to support bare-metal devices



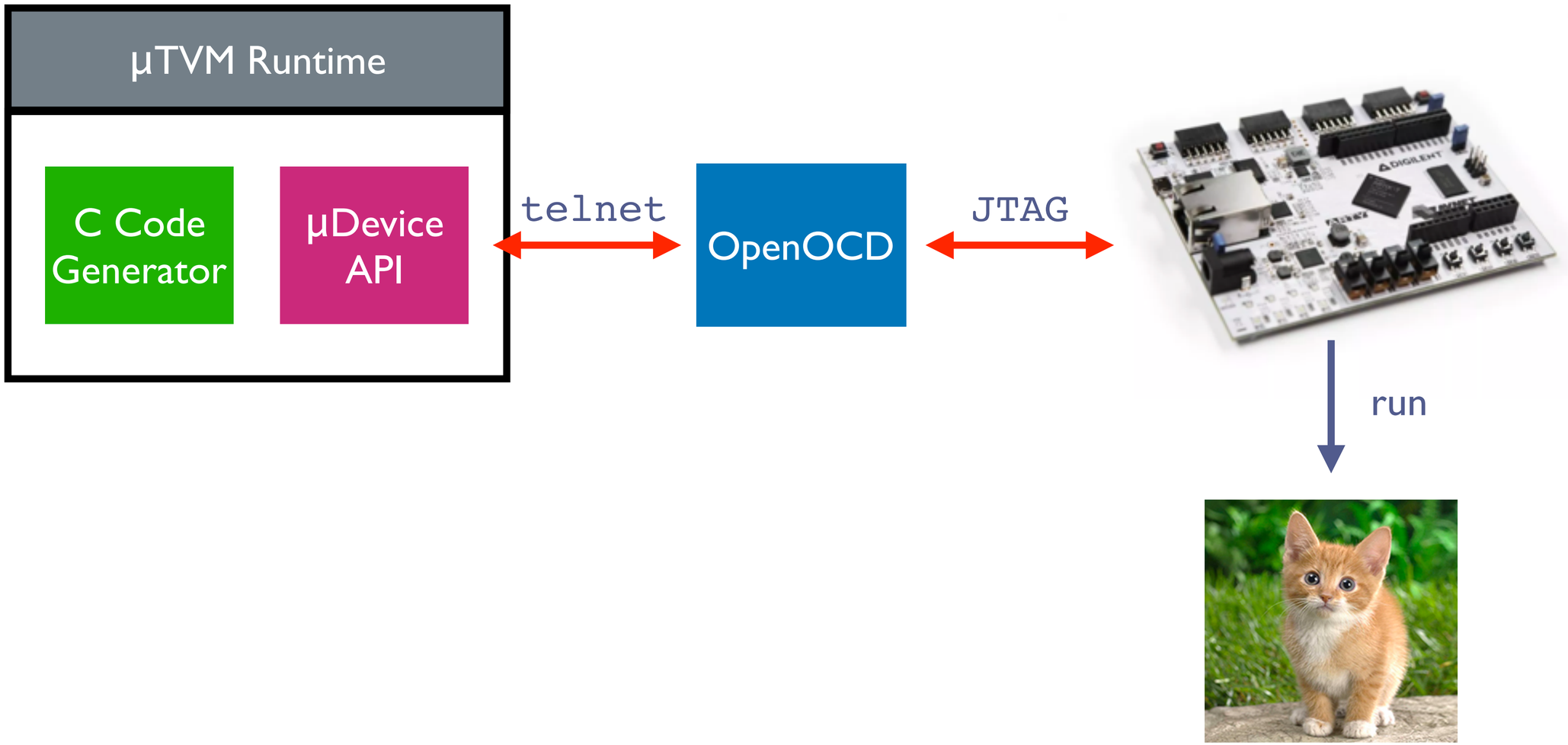
Extend TVM to support bare-metal devices



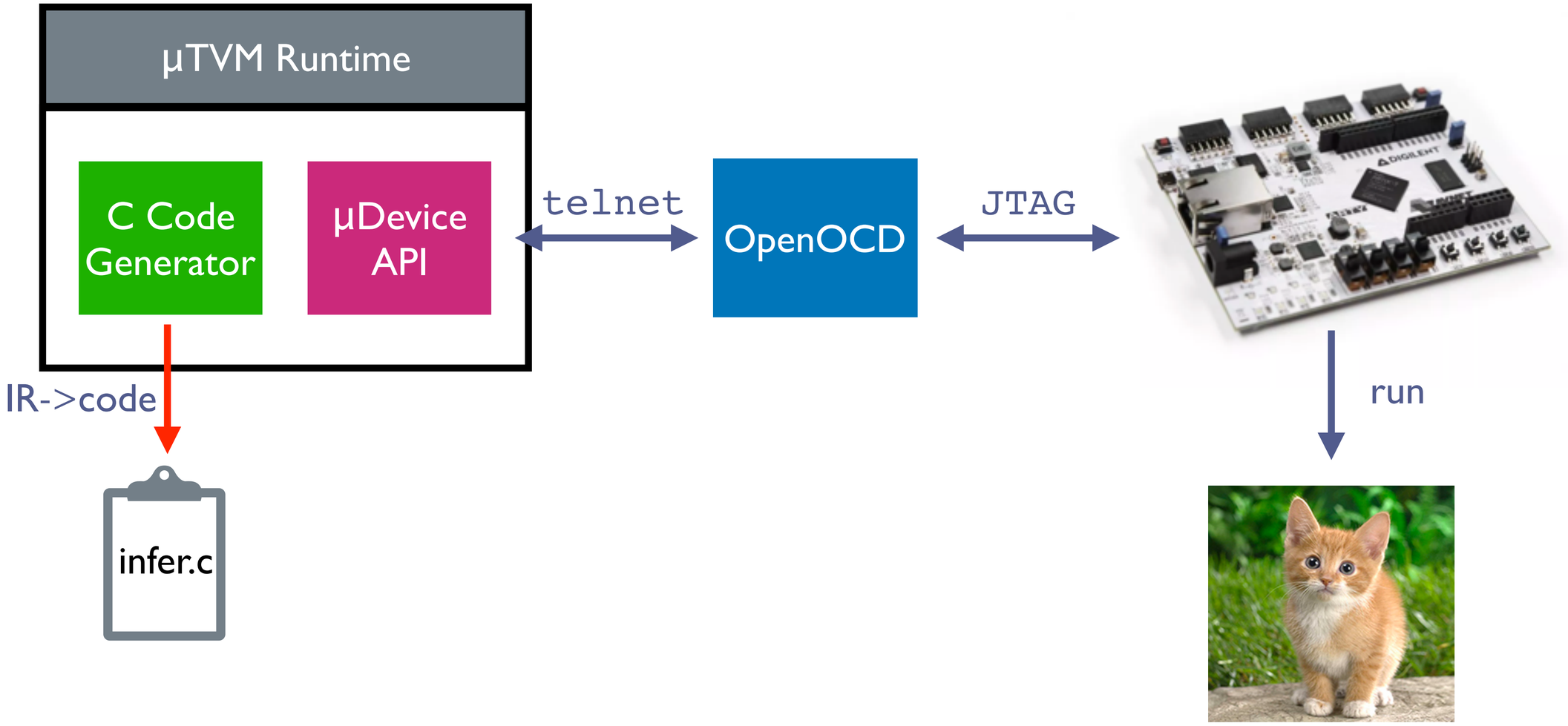
µTVM builds upon AutoTVM



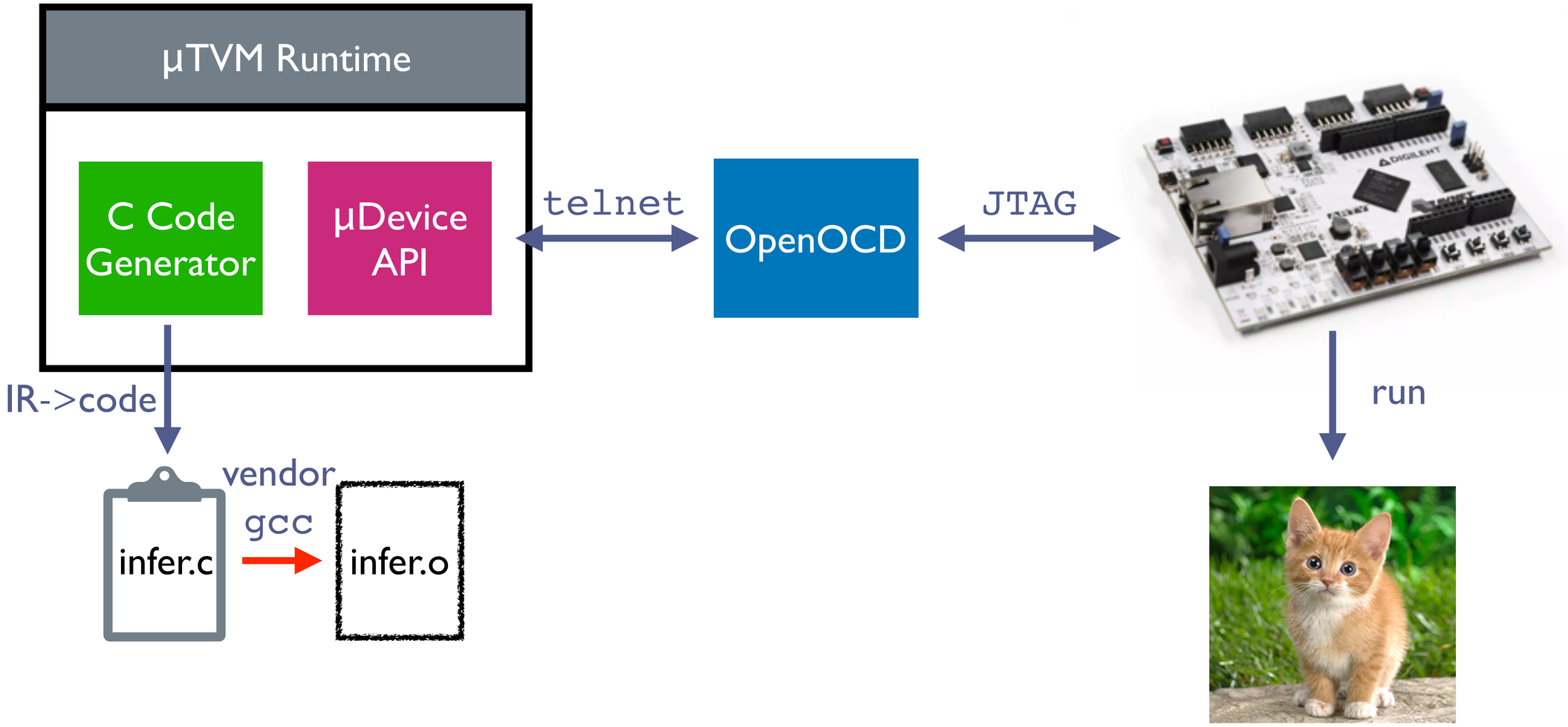
A closer look at µTVM



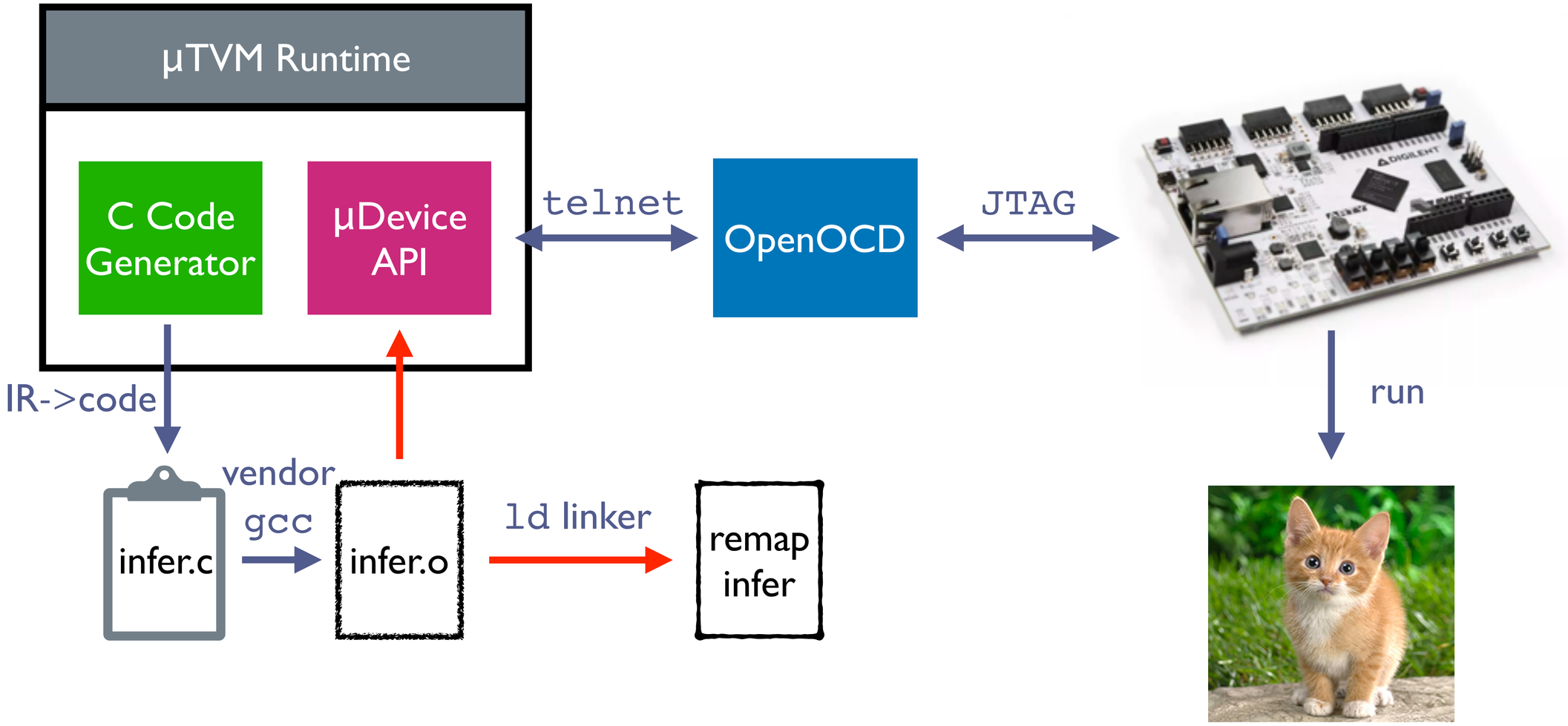
A closer look at µTVM



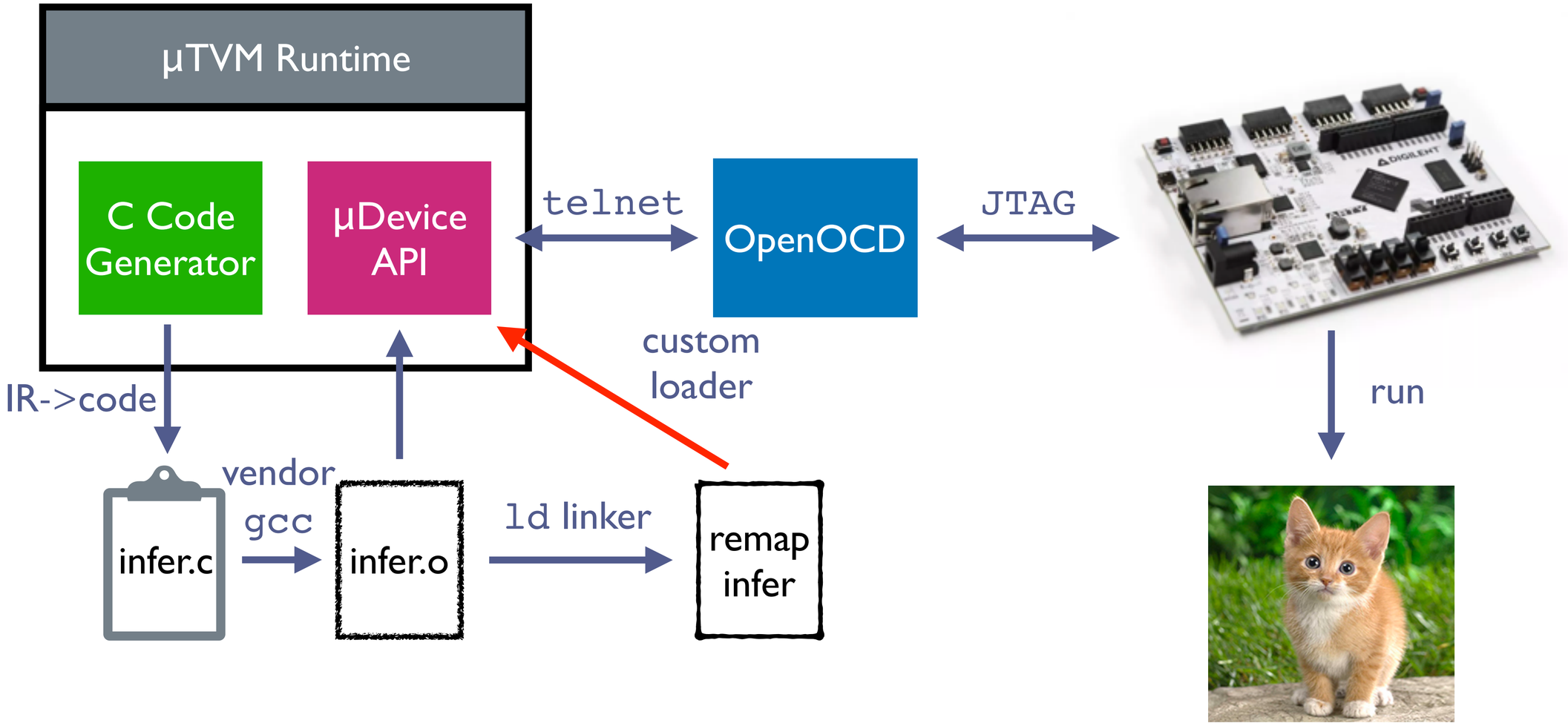
A closer look at µTVM



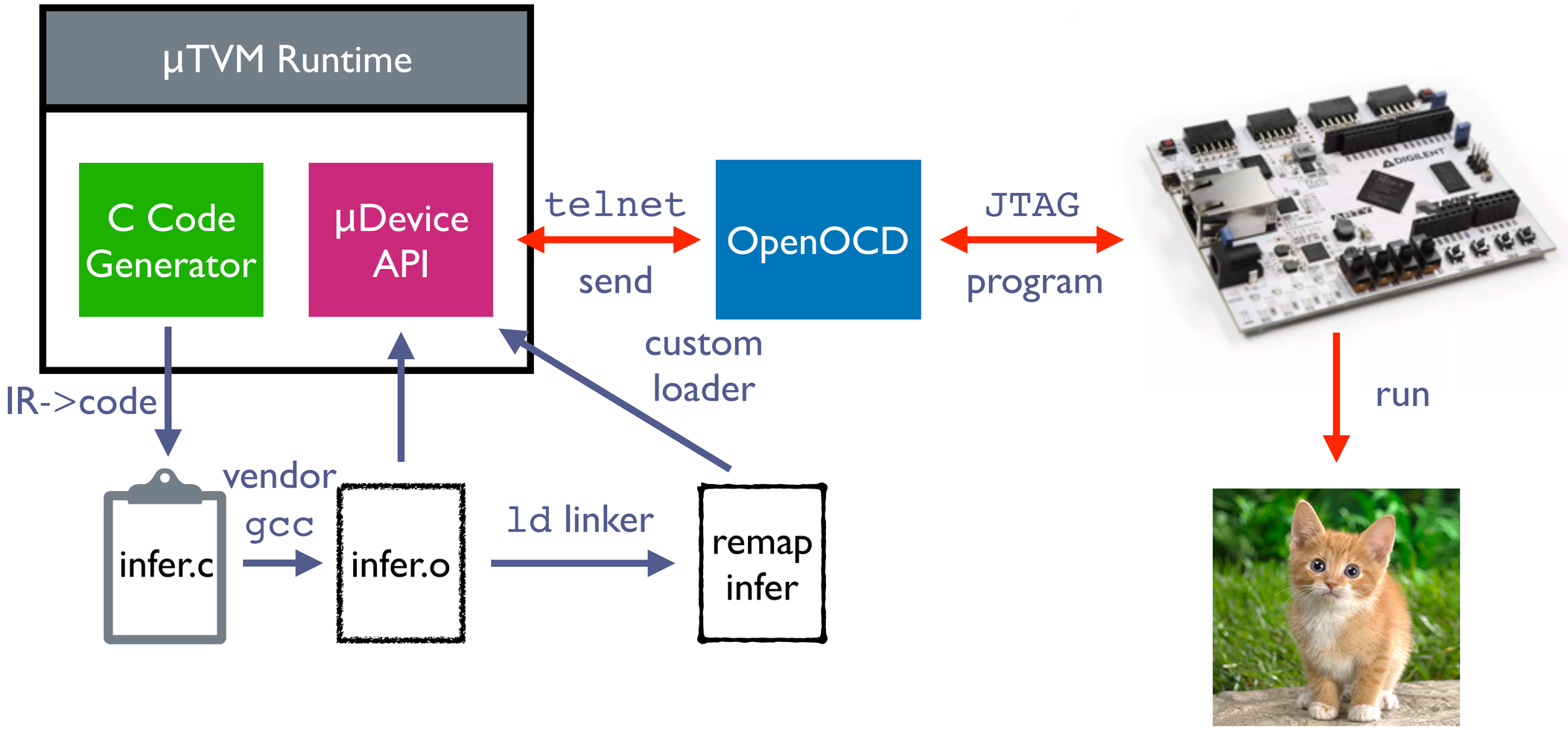
A closer look at µTVM



A closer look at µTVM

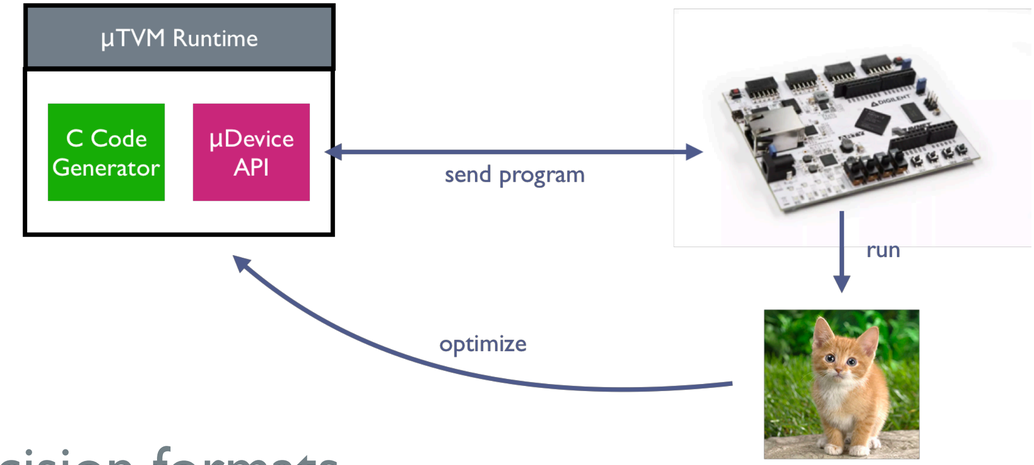


A closer look at µTVM



Next steps

- Iron out interfaces with actual hardware.
- Optimize with AutoTVM.
- Support restricted and configurable model sizes.
- Enable custom data types such as fixed-point precision formats.



Get in touch!

Pratyush Patel — patelp1@cs.uw.edu