

Choose the Right Hardware

Scenario 1: Manufacturing

Client Requirements and Potential Hardware Solution

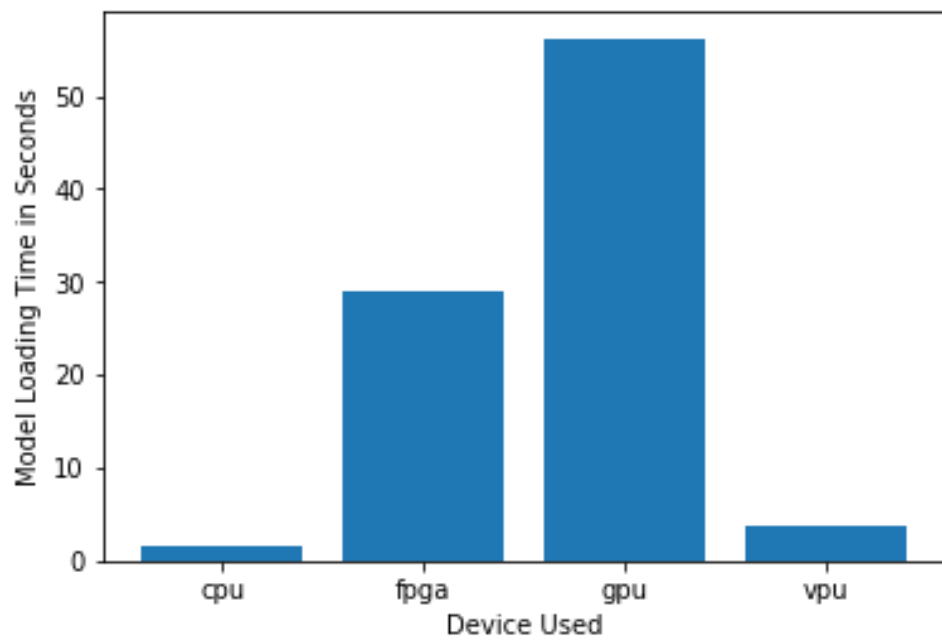
Which hardware might be most appropriate for this scenario? (CPU / IGPU / VPU / FPGA)
FPGA

Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
The system would also need to be flexible so that it can be reprogrammed and optimized to quickly detect flaws in different chip designs	FPFA can be reprogrammed and flexible enough to work in different scenario
Naomi Semiconductors has plenty of revenue to install a quality system	Any Hardware will be fine
They would ideally like it to last for at least 5-10 years.	FPGA has longer lifetime

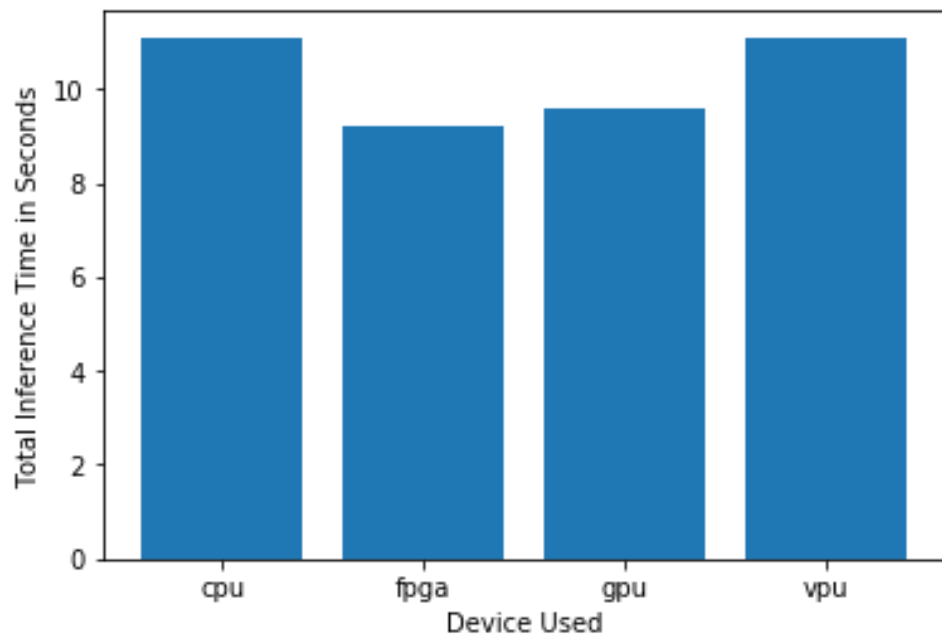
Queue Monitoring Requirements

Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or Int8)	FP16

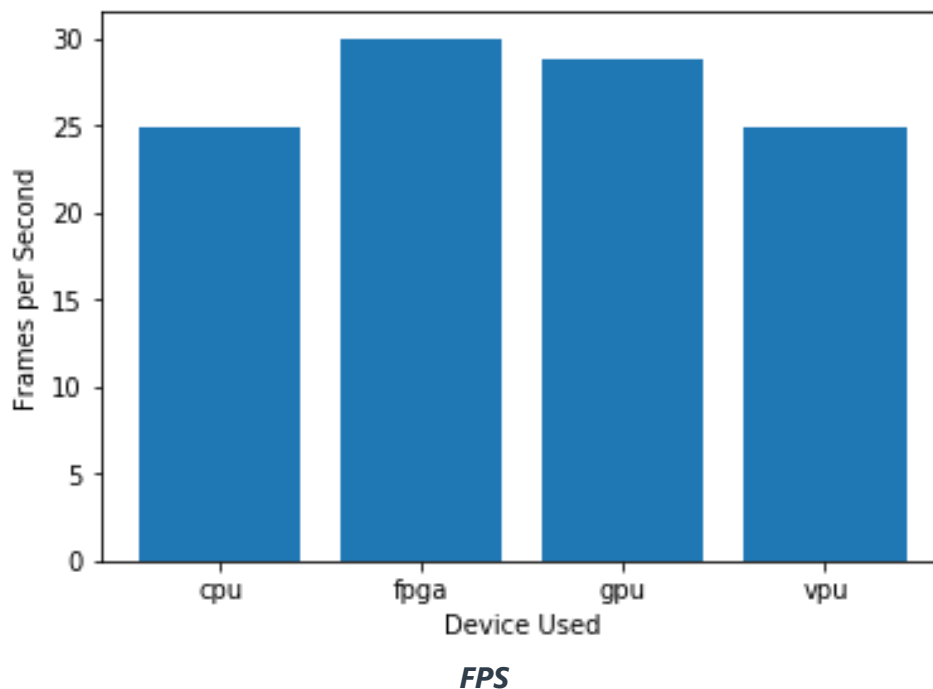
Test Results



Model Load Time



Inference Time



Final Hardware Recommendation

Write-up: Final Hardware Recommendation

- *FPGA got Best FPS and lowest Inference Time, which is well suited for industrial Need, It has comparatively higher model loading time but it will not be a problem because it is one time process.*
- *Since FPGA are more efficient (best Performance), durable and flexible for different scenarios and this project don't have budget constrain so it's the right choice*

Scenario 2: Retail

Client Requirements and Potential Hardware Solution

Which hardware might be most appropriate for this scenario?
(CPU / IGPU / VPU / FPGA)

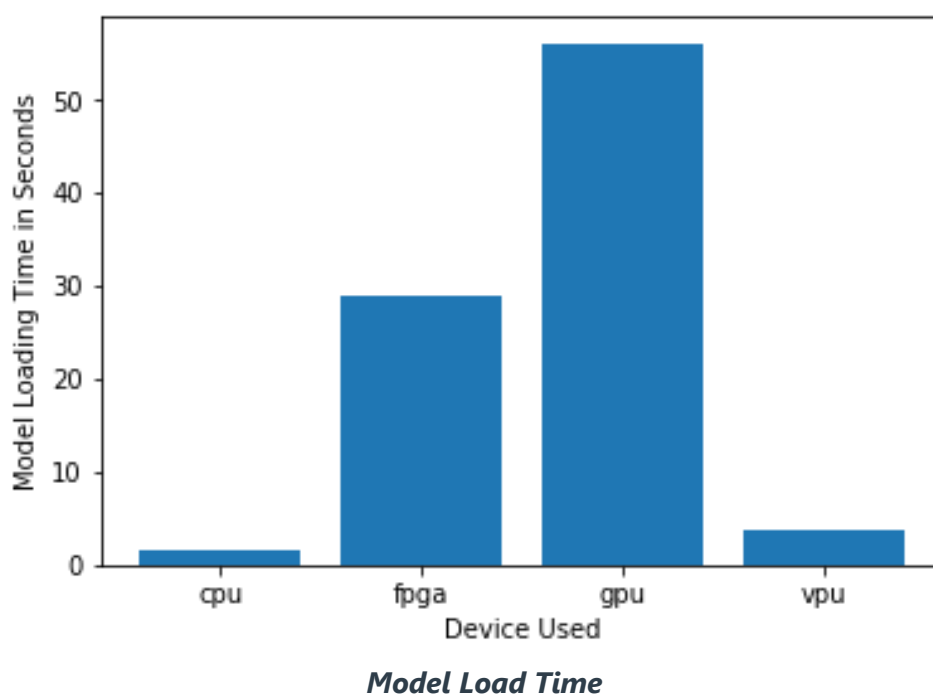
VPU

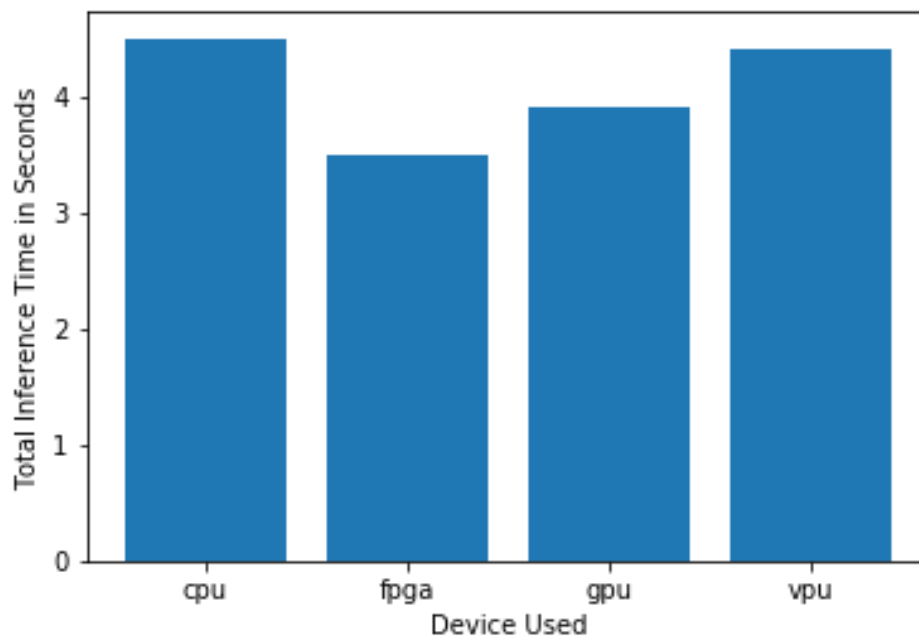
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Mr. Lin does not have much money to invest in additional hardware.	VPU or NCS2 is only cost 70-100\$, we can easily install it on multiple system. Few IGPUs are also not very costly
He would like to save as much as possible on his electric bill.	NCS2 is low-power device

Queue Monitoring Requirements

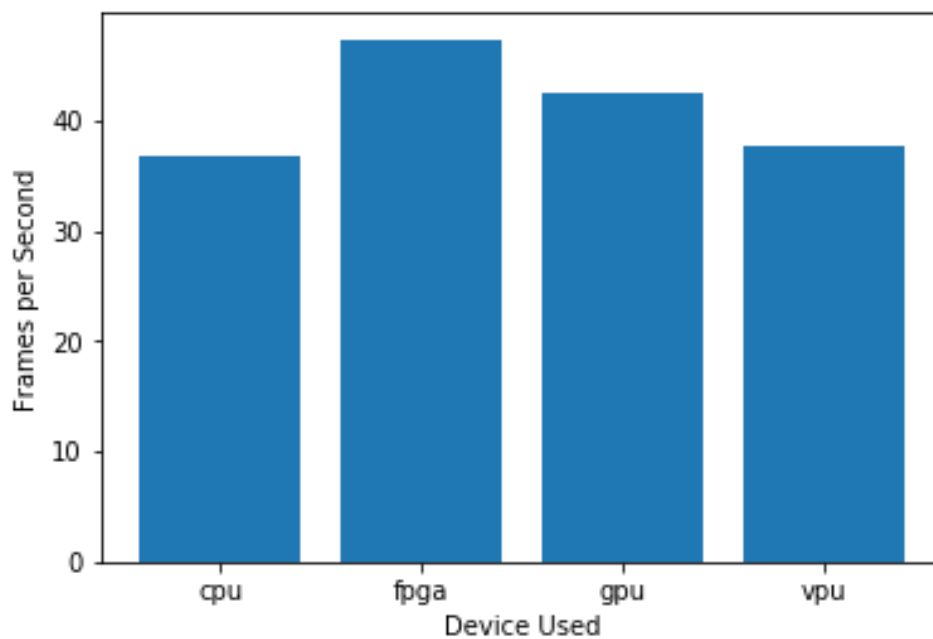
Maximum number of people in the queue	5
Model precision chosen (FP32, FP16, or Int8)	FP16

Test Results





Inference Time



FPS

Final Hardware Recommendation

Write-up: Final Hardware Recommendation

Since our client has two specific requirements, first is low cost so we cannot go for FPGA. They also need power efficient devices so NCS2 is preferable.

It also has better FPS and low inference time compare to CPU. So VPU are the ideal choice in this case

Scenario 3: Transportation

Client Requirements and Potential Hardware Solution

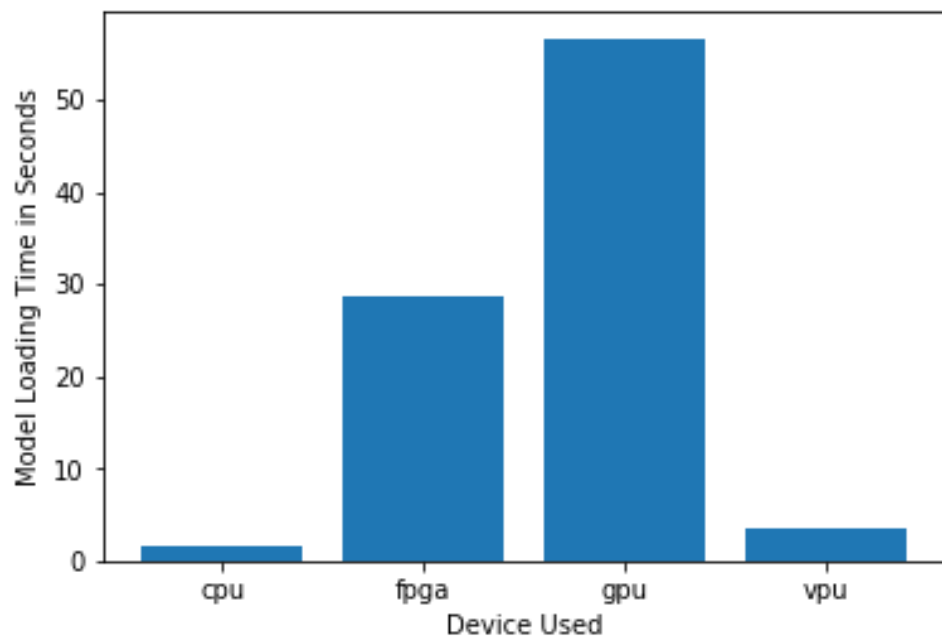
Which hardware might be most appropriate for this scenario? (CPU / IGPU / VPU / FPGA)
<i>IGPU</i>

Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Ms. Leah's budget allows for a maximum of \$300 per machine	Intel Atom E3950 IGPU which cost 57\$ or NCS2 which cost 70-100\$ can be used
They would like to save as much as possible both on hardware and future compute power requirements.	IGPU, Multiple VPUs or FPGA

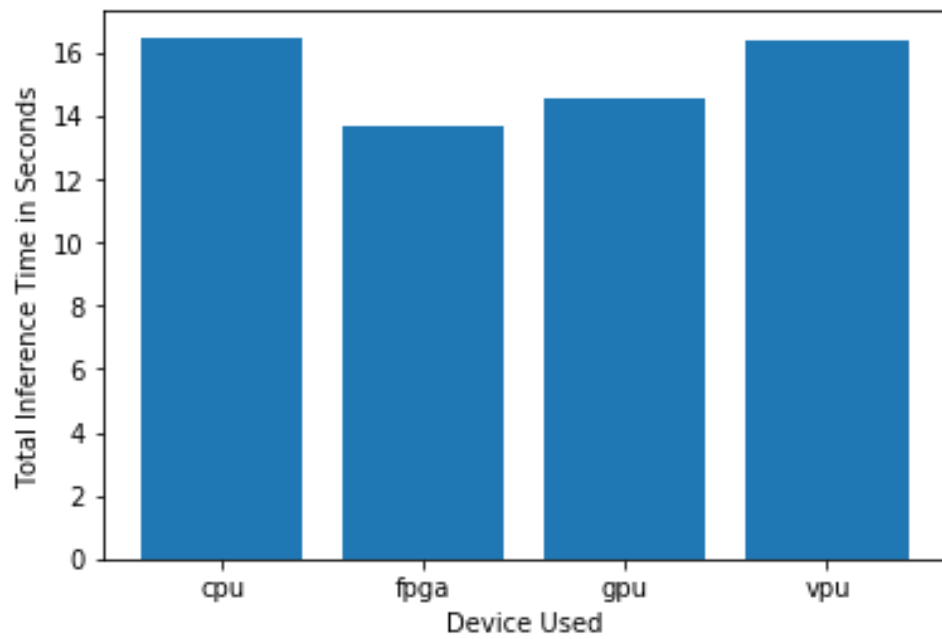
Queue Monitoring Requirements

Maximum number of people in the queue	<i>15</i>
Model precision chosen (FP32, FP16, or Int8)	<i>FP32</i>

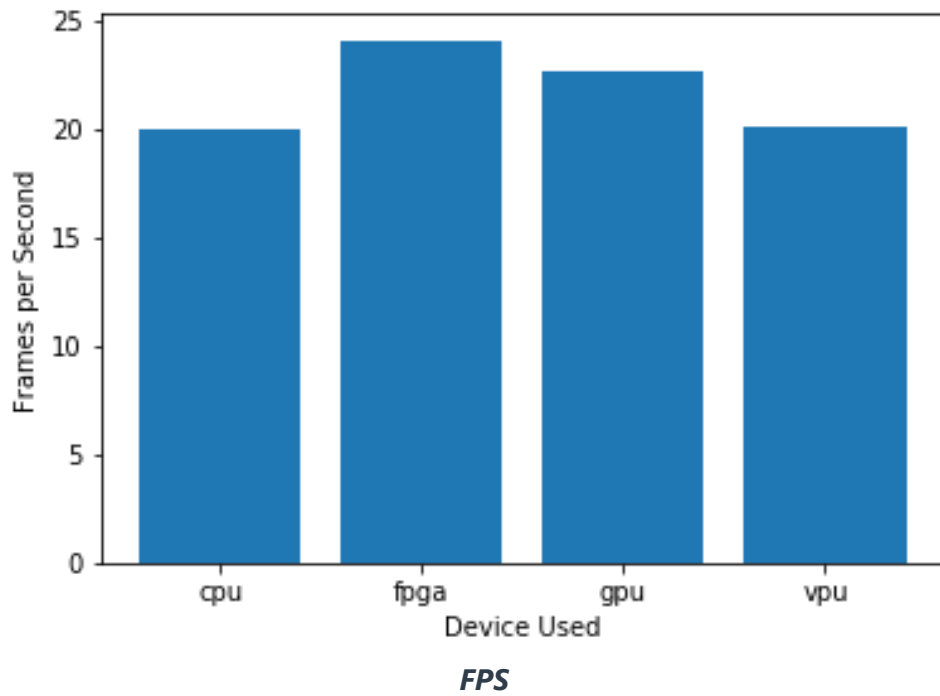
Test Results



Model Load Time



Inference Time



Final Hardware Recommendation

Write-up: Final Hardware Recommendation

*We can not use FPGA because it cost more than 300\$
Intel Atom E3950 cost around 57\$, so it satisfies our budget. So we can go for IGPU, it also has higher inference speed and FPS compare to CPUs and VPUs*