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)

FGPA

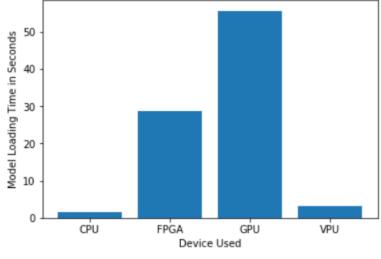
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
The client wants to run the inference in the video stream very quickly.	FGPA has low inference time which will help to attain the required performance.
The client wants a flexible system that could adapt to changing chip design.	FGPA is very flexible and hence fulfill user requirement
The client has plenty of funds to invest in this system	FGPA can be afforded by the client
The client wants the system to last for 5-10 years	FGPA has work-life for at least 10 years

Queue Monitoring Requirements

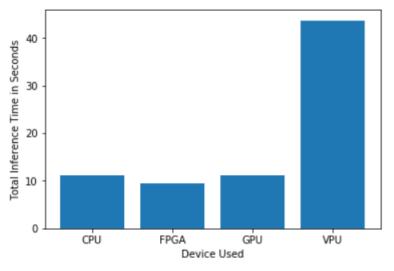
Maximum number of people in the queue	5
Model precision is chosen (FP32, FP16, or Int8)	FP16

Test Results

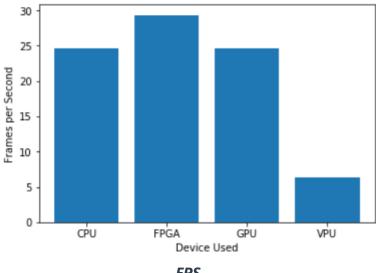




Model Load Time



Inference Time



FPS

Final Hardware Recommendation

Write-up: Final Hardware Recommendation

I have chosen FPGA because it has the highest frames per second, at least Inference time which are crucial for a high-performance system. FPGA also fulfills the client's need for a flexible system. It has a long Worklife of 10 years which is required by the client.

Scenario 2: Retail

Client Requirements and Potential Hardware Solution

Look through the scenario and find any relevant client requirements. Then, suggest a potential hardware type and explain how this hardware would satisfy each of the requirements.

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

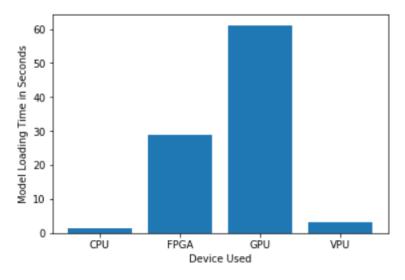
CPU

Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
The client doesn't want to buy any additional hardware	Intel i7 processor installed in the computers at checkout counters can be utilized
The client wants to use power-efficient system	Intel i7 is pretty power-efficient

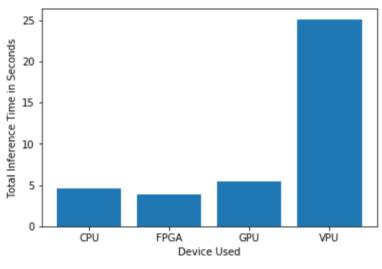
Queue Monitoring Requirements

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

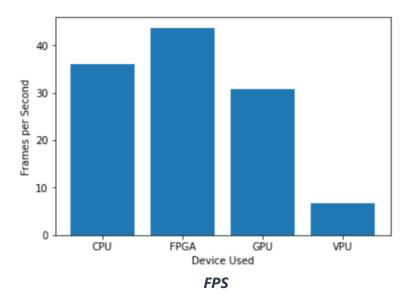




Model Load Time



Inference Time



Final Hardware Recommendation

Write-up: Final Hardware Recommendation

After analyzing the results of all hardware I recommend CPU because it's it has pretty good performance without buying any additional hardware. It's model loading time is lowest, inference time is 5 sec, and frames per second achieved is about 35fps which is pretty good. Power consumption of CPU is lower than other high-performance devices, hence CPU fulfills all the requirements of the client.

Scenario 3: Transportation

Client Requirements and Potential Hardware Solution

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

Requirement Observed	How does the chosen hardware meet this requirement?
(Include at least two.)	

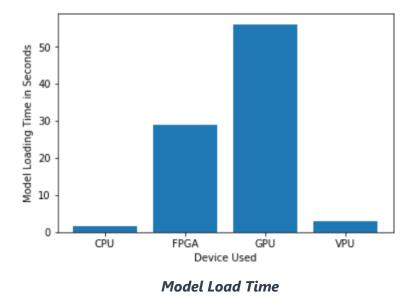


The client requires a tiny device to be connected to their CPU and the client's maximum budget is 300\$ per computer.	VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
The client wants a power-efficient device	VPU is very power efficient with Power Consumption ~1W
The client has All-in-One pc	NCS2 supports USB plug and play so it can easily connect with the client's All-in-One pc

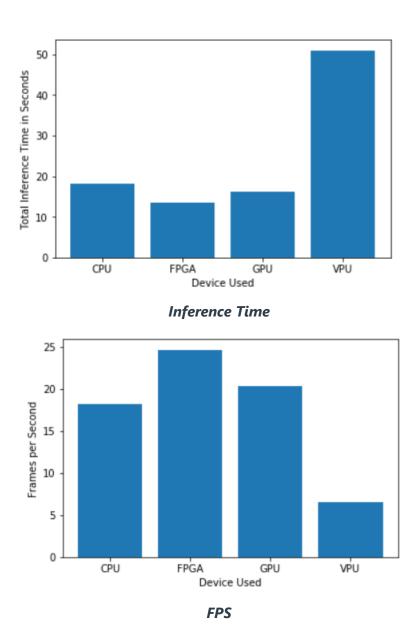
Queue Monitoring Requirements

Maximum number of people in the queue	15
Model precision is chosen (FP32, FP16, or Int8)	FP16

Test Results







Final Hardware Recommendation

Write-up: Final Hardware Recommendation

After analyzing all the results I recommend VPU(Neural Compute Stick 2) because it is very power efficient, comes in a USB form factor that can easily plugin to the client's All-in-One pc, it's cost comes under the client's budget and its performance is also sufficient to fulfill the requirements of the client.

