Choose the Right Hardware

Proposal Template

Scenario 1: Manufacturing

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)

FPGA

Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Durability: "Workers alternate shifts to keep the floor running 24 hours a day so that packaging continues nonstop"	FPGA is designed to have 100% on-time performance (24/7/365)
Performance: "To be able to detect chip flaws without slowing down the packaging process, the system would need to be able to run inference on the video stream very quickly"	FPGAs can execute neural networks with high performance and very little latency. The high performance comes from the ability to run many sections of the FPGA in parallel.
Flexibility: "because there are multiple chip designs—and new designs are created regularly—the system would also need to be flexible so that it can be reprogrammed and optimized to quickly detect flaws in different chip designs"	FPGA can be reprogrammed to adapt to new, evolving, and custom networks.

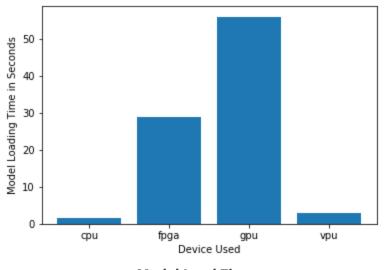
Queue Monitoring Requirements

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

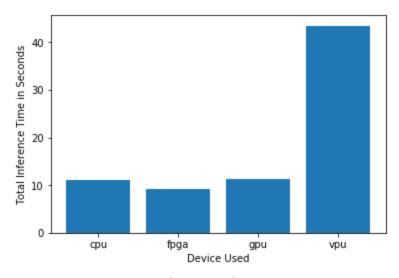
Test Results



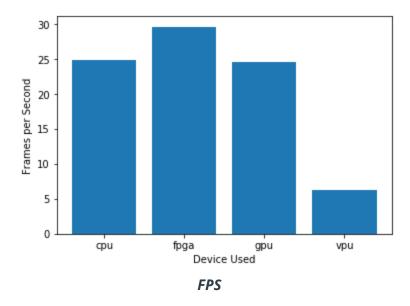
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



Model Load Time



Inference Time



Final Hardware Recommendation

Now synthesize your points from above and provide a brief write-up describing why the chosen hardware is the best choice for this scenario. Be sure to discuss the client's requirements, the test results, and how these relate to one another (e.g., perhaps one of the devices performed better than the rest, but does not meet one of the client's requirements).

Write-up: Final Hardware Recommendation

FPGA is showing the best performance in comparison with other devices, and also meets requirements for durability and flexibility.

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?
Cost: "Mr. Lin does not have much money to invest in additional hardware"	"Most of the store's checkout counters already have a modern computer, each of which has an Intel i7 core



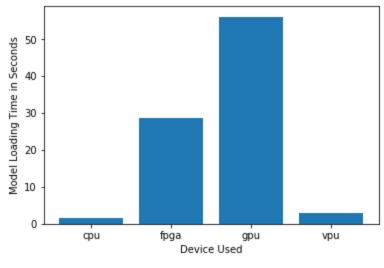
	<i>processor"</i> . So no additional hardware costs are required.
Power consumption: "Mr. Lin would like to save as much as possible on his electric bill"	Intel Core i7 CPUs are devices with relatively low power consumption.

Queue Monitoring Requirements

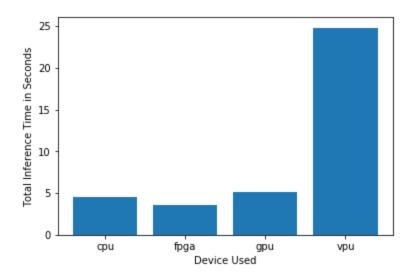
Maximum number of people in the queue	3
Model precision chosen (FP32, FP16, or Int8)	FP32

Test Results

After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).

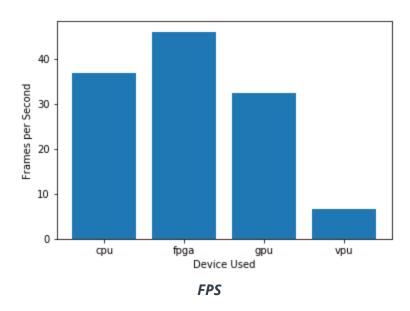


Model Load Time





Inference Time



Final Hardware Recommendation

Now synthesize your points from above and provide a brief write-up describing why the chosen hardware is the best choice for this scenario. Be sure to discuss the client's requirements, the test results, and how these relate to one another (e.g., perhaps one of the devices performed better than the rest, but does not meet one of the client's requirements).

Write-up: Final Hardware Recommendation

CPU performance is good enough, and CPU is the best choice for application because no additional hardware costs are required.

Scenario 3: Transportation

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 appro (CPU / IGPU / VPU /	
VPU	

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



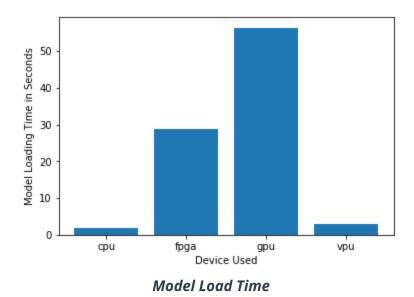
Cost: "Ms. Leah's budget allows for a maximum of \$300 per machine"	Intel NCS2 is a low cost device with a price of \$100
Power consumption: "The CPUs in these machines are currently being used to process and view CCTV footage for security purposes and no significant additional processing power is available to run inference".	Intel NCS2 power consumption is ~1W
Size: "They monitor the entire situation with 7 CCTV cameras on the platform. These are connected to closed All-In-One PCs that are located in a nearby security booth".	Intel NCS2 has a size of a thumb drive, can be easily plugged into a USB port.

Queue Monitoring Requirements

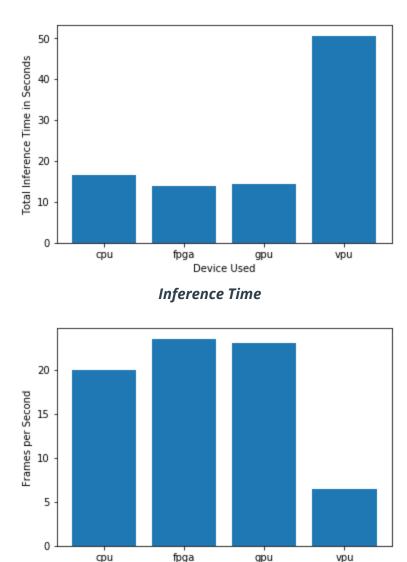
Maximum number of people in the queue	7
Model precision chosen (FP32, FP16, or Int8)	FP16

Test Results

After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).







Final Hardware Recommendation

Now synthesize your points from above and provide a brief write-up describing why the chosen hardware is the best choice for this scenario. Be sure to discuss the client's requirements, the test results, and how these relate to one another (e.g., perhaps one of the devices performed better than the rest, but does not meet one of the client's requirements).

Device Used

FPS

Write-up: Final Hardware Recommendation

VPU performance is very low in comparison with other devices, but with Multi-Device Plugin iit's possible to distribute inference tasks between multiple NCS sticks plugged into one machine (since \$300 budget is confirmed, 3×100 NCS sticks per machine would be possible).

