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.

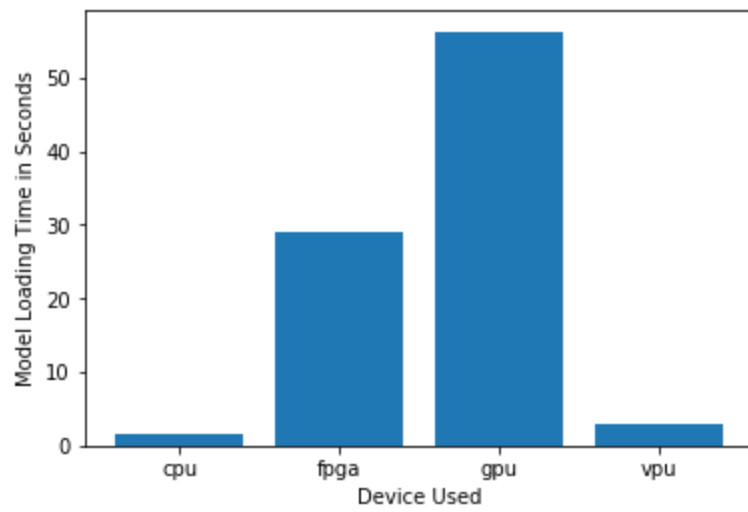| Which hardware might be most appropriate for this scenario? (CPU / IGPU / VPU / FPGA) |
|---|
| *FPGA* |

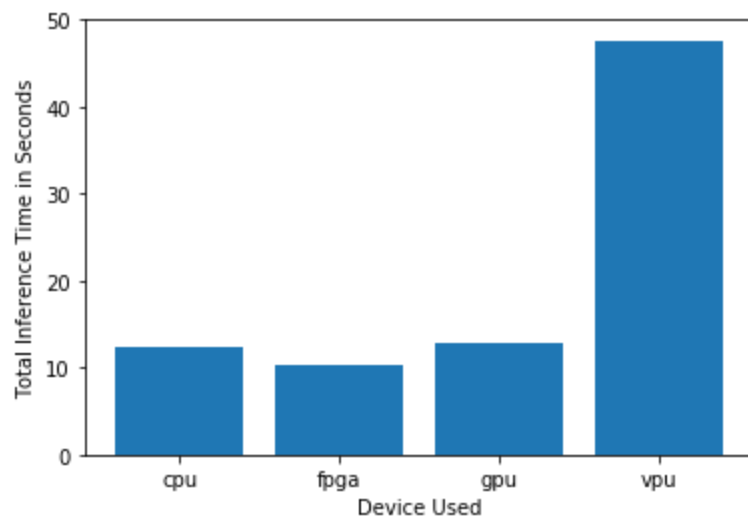| Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Example requirement: The client requires a tiny device to be connected to their CPU—and their budget is only about $100 for each device.	*Example explanation:* VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
The life span needs to be somewhere between 5-10 years. That means it needs to be durable.	FPGA's have long life spans. The devices from Intel IoT group can last up to 10 years.
To address the second issue the client wants to repurpose the device. Which means it needs to be flexible.	FPGA's are reprogrammable and can be updated for different requirements without changing the hardware.

Queue Monitoring Requirements

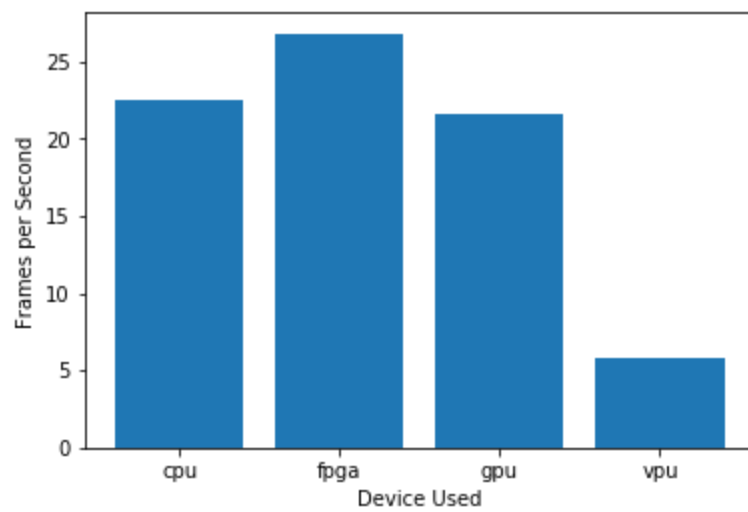|	
Maximum number of people in the queue	3
Model precision chosen (FP32, FP16, or Int8)	FP32

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

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).

Write-up: Final Hardware Recommendation
FPGA is durable and reprogrammable which are both in client's demands as the client wants the device to last for 5-10 years and also the client wants it to be flexible to address the issue of flawed chips.   Also when it comes to inference time and FPS the FPGA performs better than every other available hardware. When it comes to model load time although CPU and VPU are better but durability and cost factors come into play when CPU is considered and in case of VPU the inference time is way too high considering this scenario.

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.

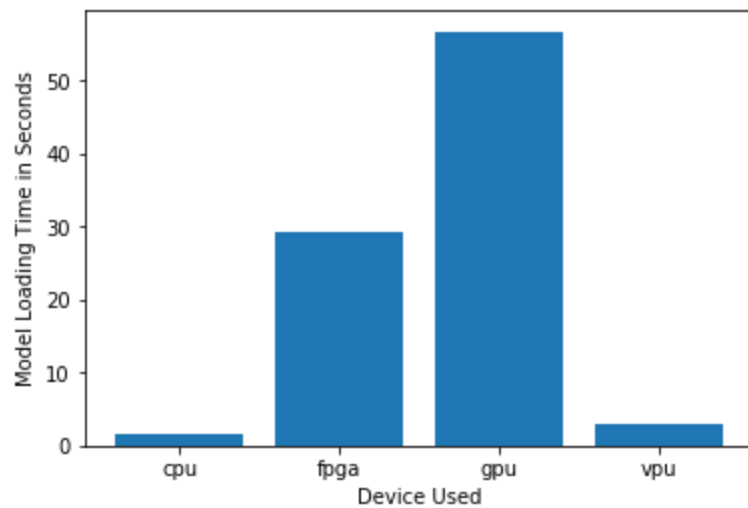| Which hardware might be most appropriate for this scenario?
 (CPU / IGPU / VPU / FPGA) |
|---|
| IGPU |

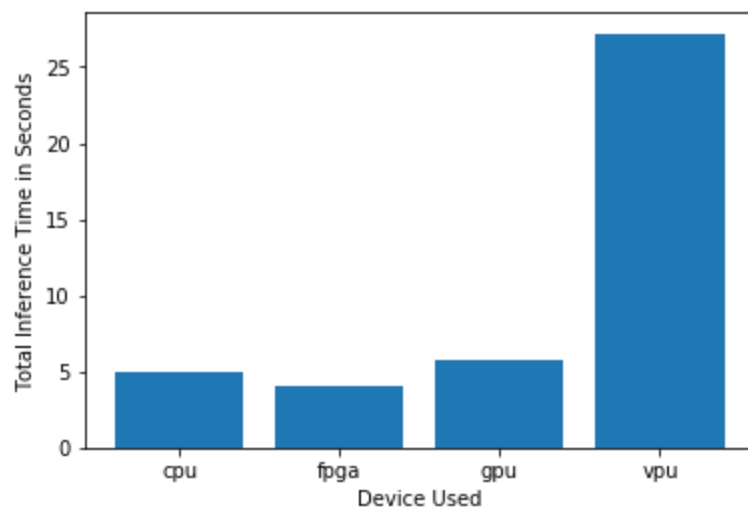| Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Example requirement: The client requires a tiny device to be connected to their CPU—and their budget is only about $100 for each device.	*Example explanation:* VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
Client do not want to invest much	IGPU are integrated in CPUs so we just need to configure the existing CPUs for the operations
Need to save electricity	It can easily be powered down to save electricity when there are less people at the store.

Queue Monitoring Requirements

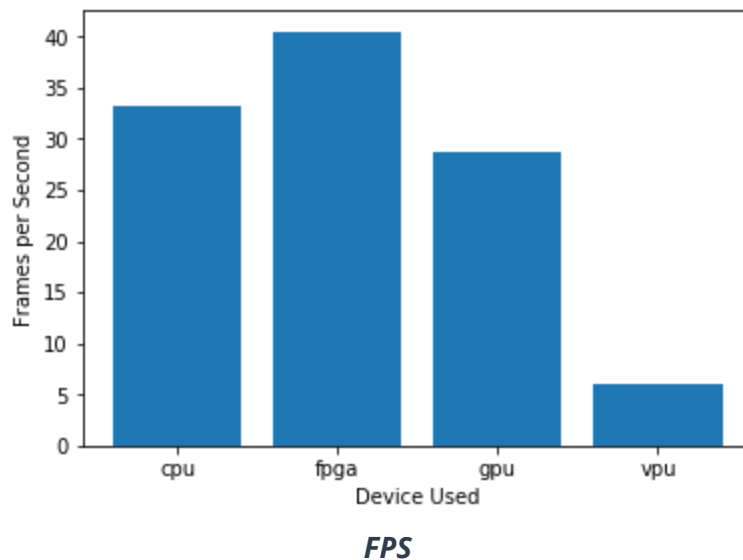| Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or Int8)	FP32

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

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).

Write-up: Final Hardware Recommendation
IGPU is recommended because in this case the client has good CPUs already and it will be really cost effective to go with IGPU. The FPS processing is comparable to CPU and FPGA. Same goes for the total inference time. Although the model loading time is high for IGPU but that needs to be done only once before starting for the day so that's acceptable.

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 appropriate for this scenario? (CPU / IGPU / VPU / FPGA)
VPU

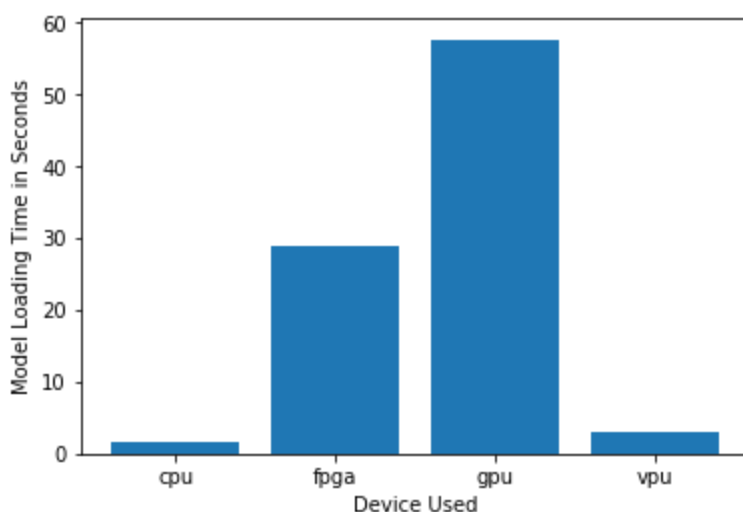| Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?

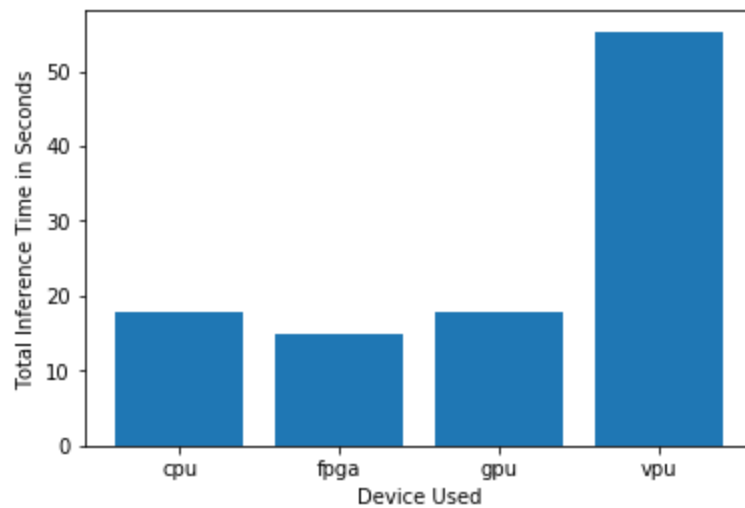| Example requirement:	Example explanation:
The client requires a tiny device to be connected to their CPU—and their budget is only about $100 for each device.	VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
Budget of around 300$	VPU are cheaper compared to other hardware
No additional processing power present	VPU is a low power device

Queue Monitoring Requirements

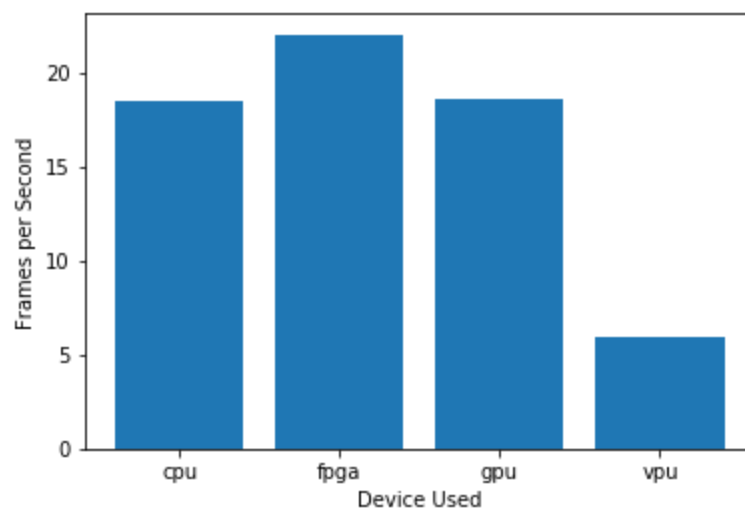| Maximum number of people in the queue	5
Model precision chosen (FP32, FP16, or Int8)	FP16

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

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).

Write-up: Final Hardware Recommendation
Due to cost and power reasons VPU is recommended. Clearly there's a limit of processing power because of 7 CCTVs and no significant power left to run inference. The model loading time is really low which is better than FPGA and GPU. Although the inference time is higher and the FPS is low but this tradeoff is acceptable considering the power issue.