

Choose the Right Hardware

Proposal

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

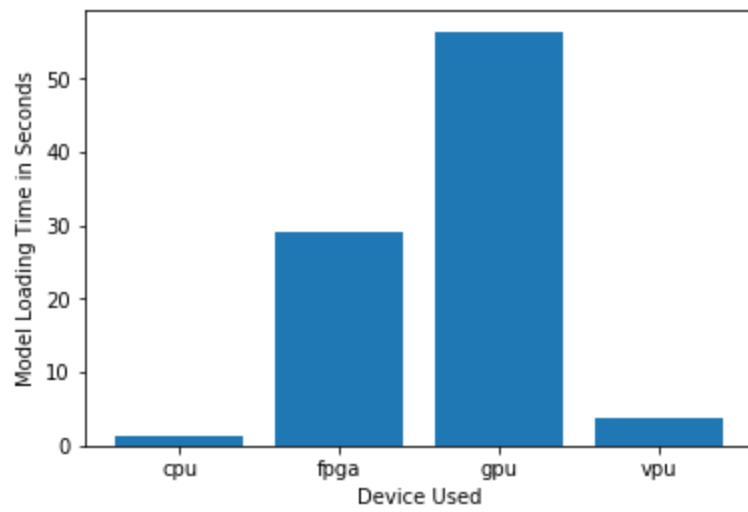
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
<i>Money is not an issue for the customer</i>	<i>FPGA is the best choice because of how versatile it is.</i>
<i>The client needs at least 5 image processing tasks to be completed per second which will require good performance</i>	<i>FPGA can certainly handle this. Due to their high performance and low latency this can be made possible.</i>
<i>The client also needs to be able to detect flaws in new chip models that will be released in the future.</i>	<i>Since the FPGAs can be reprogrammed, it is possible to meet this requirement.</i>

Queue Monitoring Requirements

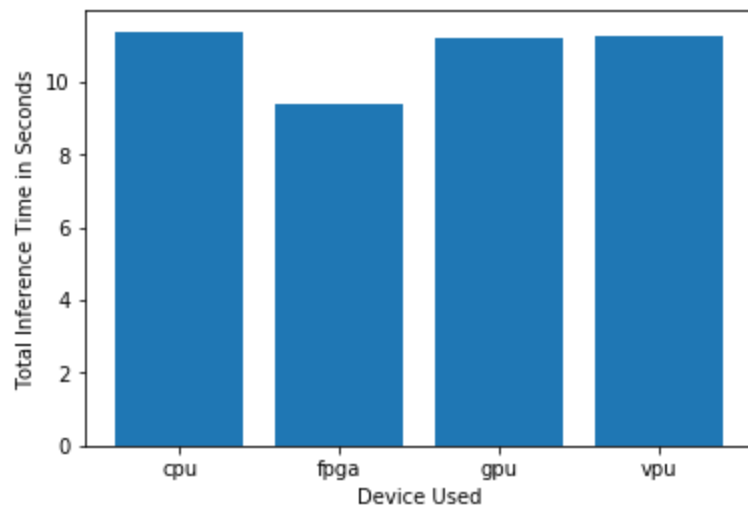
Maximum number of people in the queue	4
Model precision chosen (FP32, FP16, or Int8)	FP32

Test Results

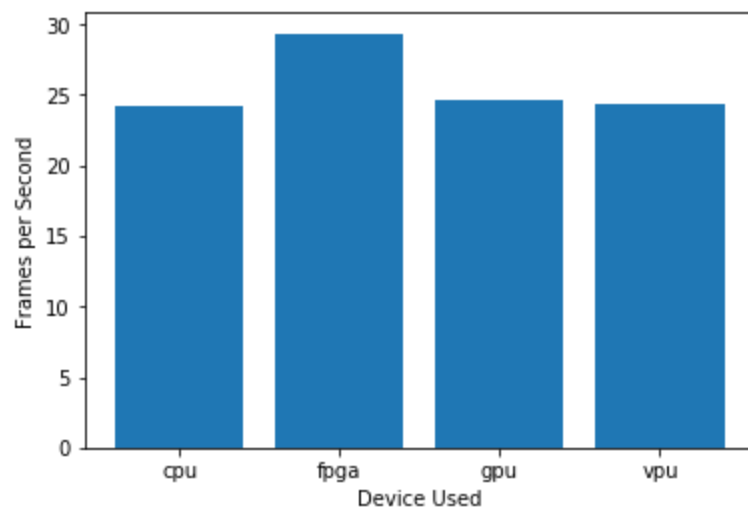
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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

Due to their high performance-low latency, ability to be reprogrammed, ability to last over 10+ years makes them the best choice for this 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)

IGPU - This is possible the best choice for this client.

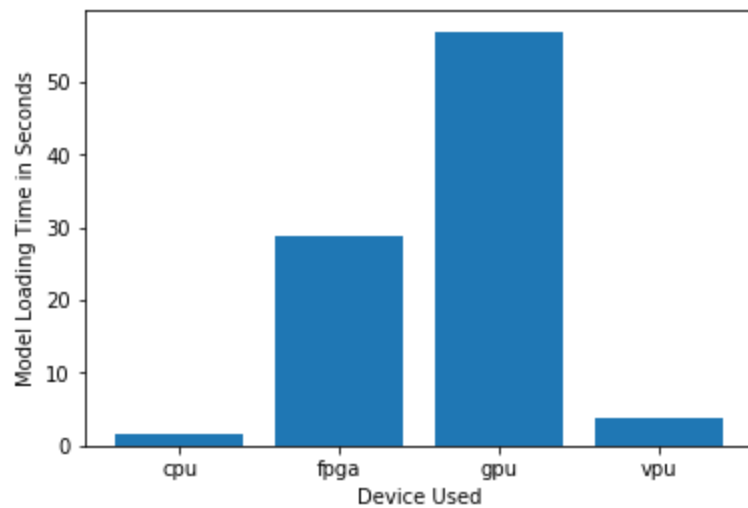
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
<i>Since the client is on a budget, upgrading is not a very exciting option.</i>	<i>The client can use the CPU which is a fairly powerful processor but it is not used to its potential.</i>
<i>Although the store has fairly good sales, the profit margin, however, is less. The manager also believes in fair employment practices, and the computers are not utilised to the full potential</i>	<i>Upgrading to newer, more powerful hardware may hamper either the profit to the company or compensation to the employees. So taking advantage of the available hardware is beneficial</i>
<i>17 CPU's use moderate amount of energy, not as high as Xeon processors, however the computers must run for prolonged period of time</i>	<i>The will help save electricity bill in the long run, and maybe not a lot, but it will improve the overall profit of the company.</i>

Queue Monitoring Requirements

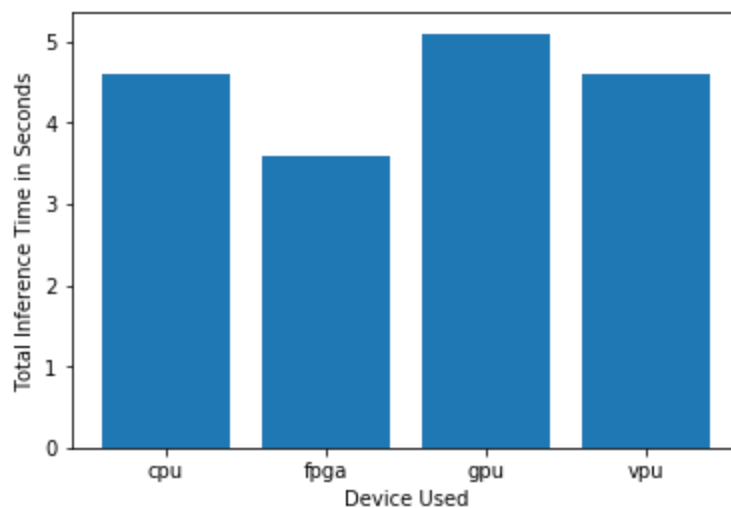
Maximum number of people in the queue	2-5
Model precision chosen (FP32, FP16, or Int8)	FP32(since, detecting people in the queue is essential)

Test Results

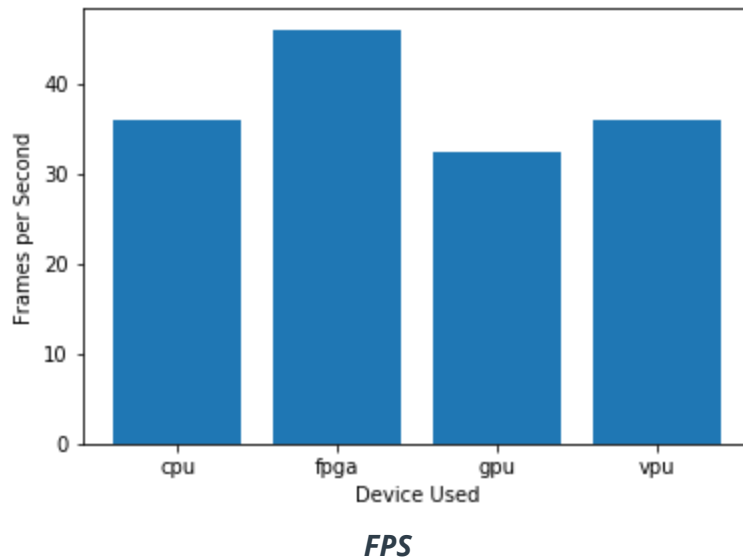
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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

Considering the company makes very marginal profit while maintaining good employment standards, it is best to use the existing CPU while making use of the IGPU for Edge AI tasks. This will improve the congestion, reduce electricity bills and increase profit since customers no longer have to wait for prolonged periods of time in the queues.

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 - Because the client wants to save both money and power, as much as possible, VPU is the best choice. However for this particular scenario FPGA might have been the best choice - good performance, 10+ years lifespan, ability to be modified, low latency, but they are costly. A CPU with IGPU would also work but for inference on a train station with a lot of people, we would need a very good CPU which would overshoot the \$300 budget.

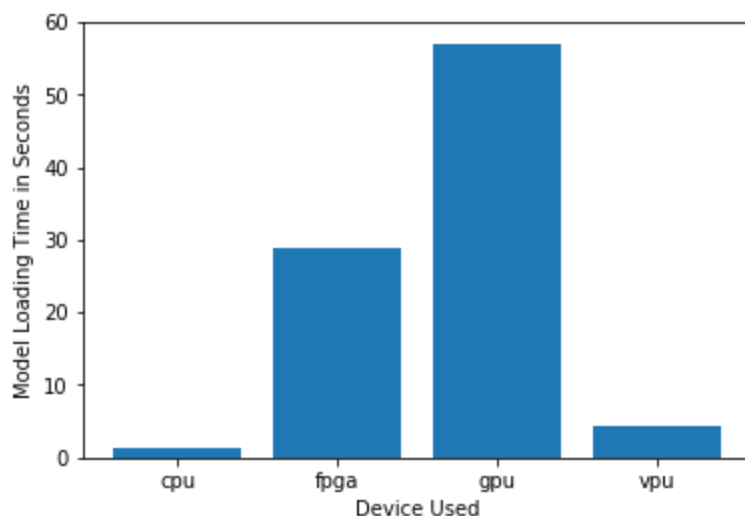
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
<i>Client wants to save both money and electricity bills</i>	<i>VPU is the best at doing it</i>
<i>The system will probably run for 24x7, 365 days, so power requirements must be kept in mind</i>	VPU satisfies this criteria. The least powerful CPU, the Intel ATOM will require about 12 times the power required by a NCS 2 stick.
<i>For improved performance we may buy upto 2-3(even more) VPU for a single system and not overshoot the \$300 budget</i>	<i>Only possible with VPU</i>

Queue Monitoring Requirements

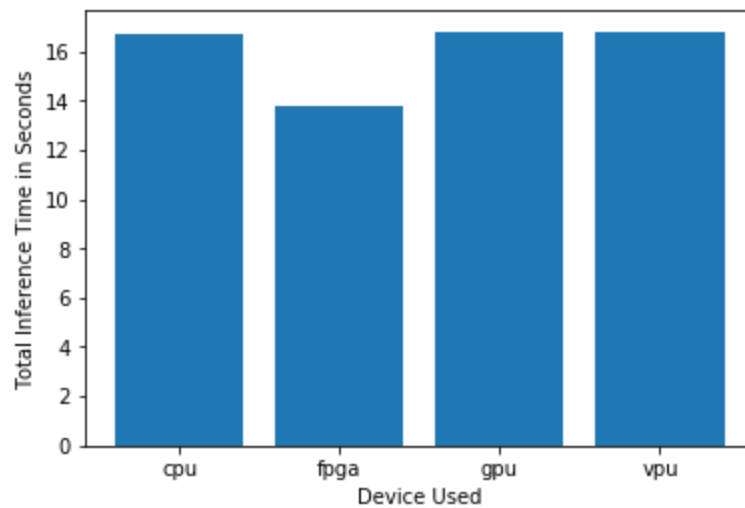
Maximum number of people in the queue	7-15
Model precision chosen (FP32, FP16, or Int8)	FP32

Test Results

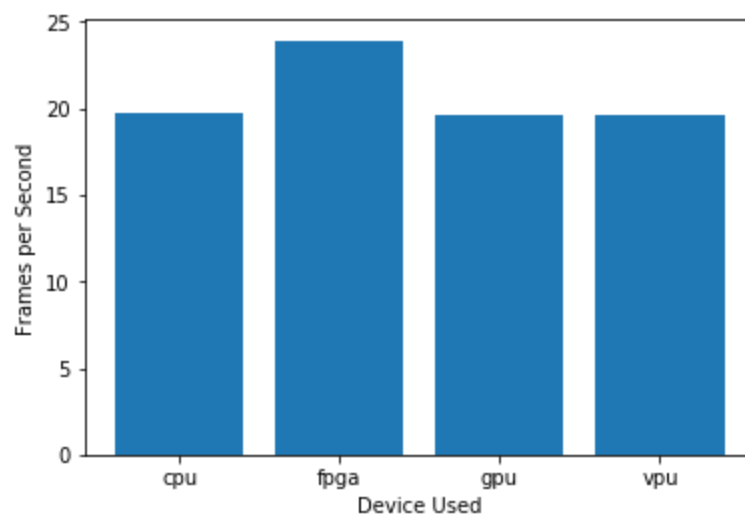
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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

Considering that the system will run non-stop for a prolonged period of time, power requirement is the first priority, also there is a \$300 cap that the customer has set, so VPU(possibly multiple VPUs per PC in parallel) is the best choice for this customer.