

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

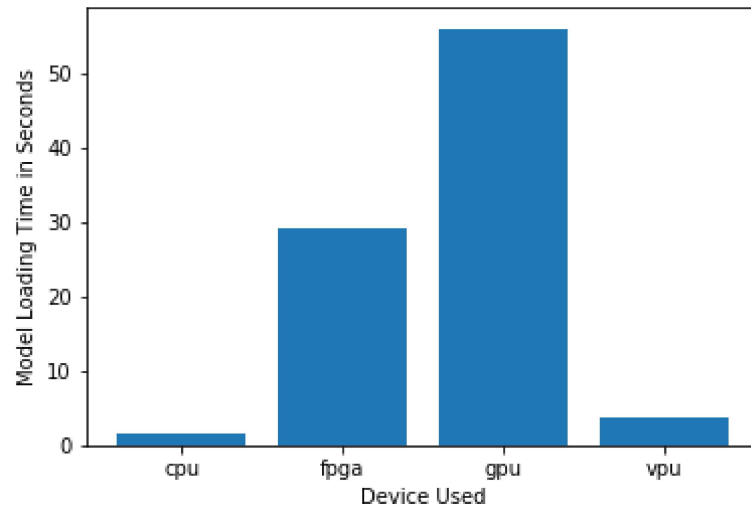
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Client initially needs hardware that can perform 5 inferences per second but need more performance later	<i>The chosen hardware FPGA have low inference time at 30 fps</i>
<i>Client would like to repurpose the hardware for detecting defects within products</i>	<i>FPGA can be repurposed after production by changing the stream and support a different model</i>
<i>Client would like the investment to last for 5-10 years</i>	<i>FPGA have guaranteed lifespan of ~10 years</i>
<i>[TODO: Type your answer here]</i>	<i>[TODO: Type your answer here]</i>

Queue Monitoring Requirements

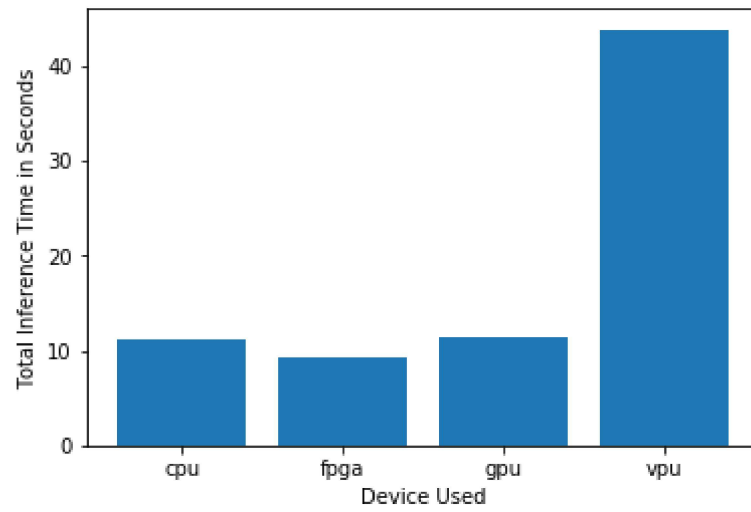
Maximum number of people in the queue	5
Model precision chosen (FP32, FP16, or Int8)	FP32

Test Results

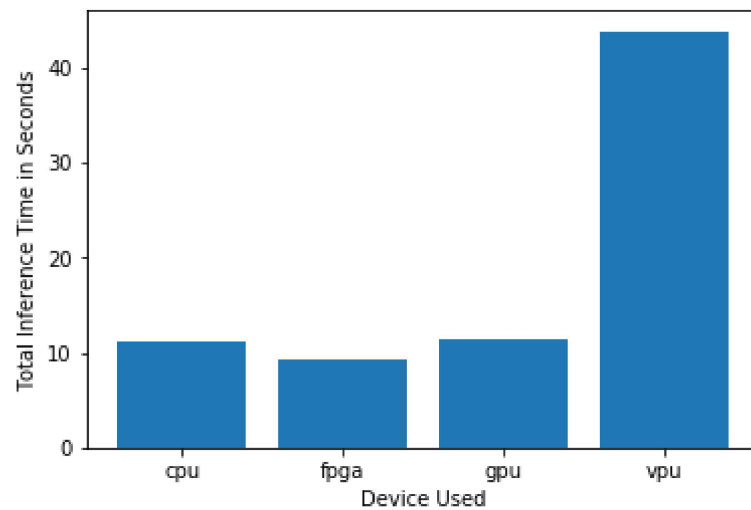
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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 would be the final hardware recommendation since the performance is important for task that would come with defect detection after the queue issues and VPU would simply not cut it. Also although the other hardware such as CPU and GPU have comparable inference time they can't be repurposed for tasks later to come.

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

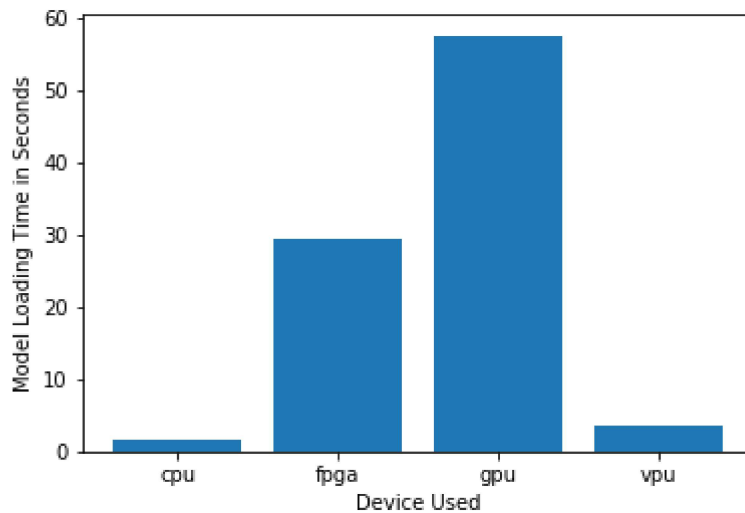
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
Client would like minimal hardware investment	Using IGPU within existing computers there would be no need for additional hardware investment
<i>Client would like low electricity costs</i>	<i>With no additional hardware the power required to support the machines should be kept low</i>
<i>[TODO: Type your answer here]</i>	<i>[TODO: Type your answer here]</i>
<i>[TODO: Type your answer here]</i>	<i>[TODO: Type your answer here]</i>

Queue Monitoring Requirements

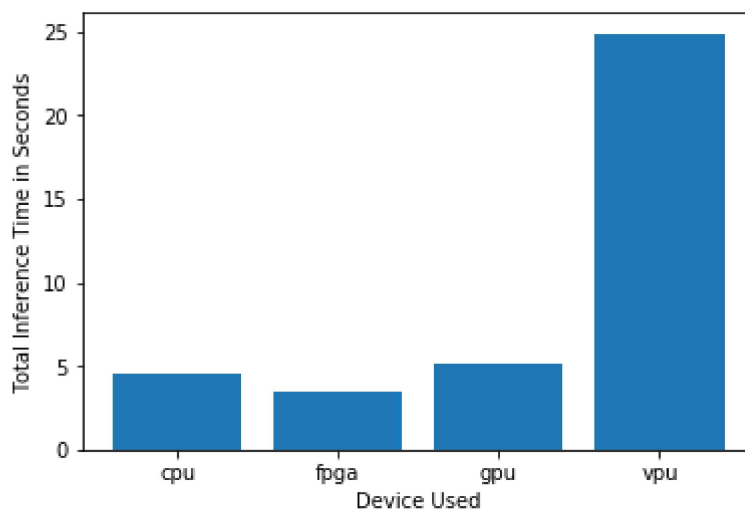
Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or	FP32

Test Results

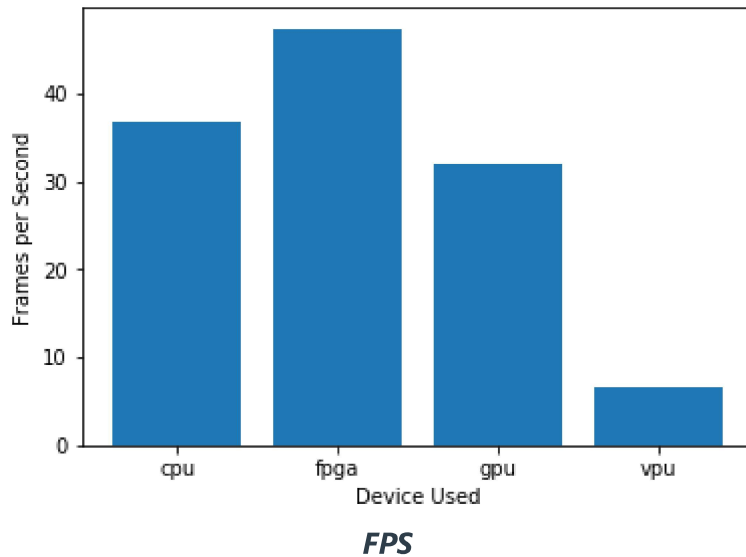
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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 after testing with the video feed it appears that CPU have comparable performance as the IGPU with respect to the video feed while having significantly lower load time. Since the store's CPUs are only doing minimal tasks it makes sense to utilize the CPU for inference as well. The FPGA and VPU would require additional hardware investment and power requirements.

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

How does the chosen hardware meet this

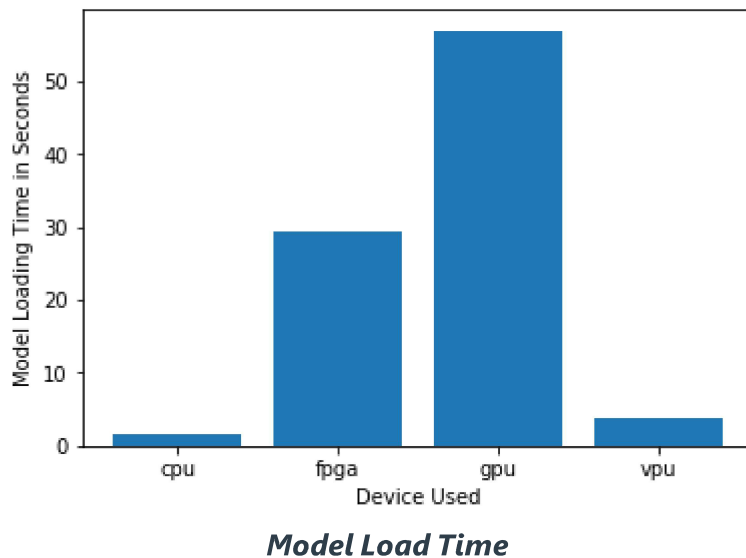
(Include at least two.)	requirement?
<i>Client have a budget of \$300 per machine</i>	VPU or NCS2 is costs under \$100 and is within the budge
<i>Client does not have additional compute for inference</i>	VPU or NCS2 can be used for the inference through the interface unit
<i>[TODO: Type your answer here]</i>	<i>[TODO: Type your answer here]</i>
<i>[TODO: Type your answer here]</i>	<i>[TODO: Type your answer here]</i>

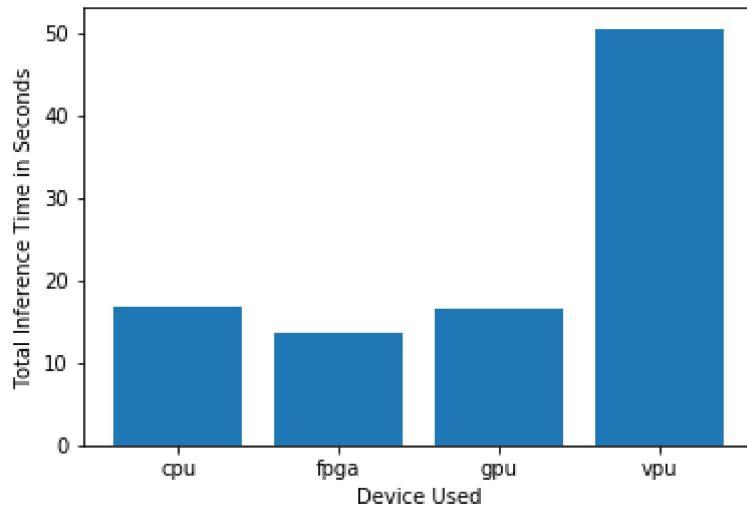
Queue Monitoring Requirements

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

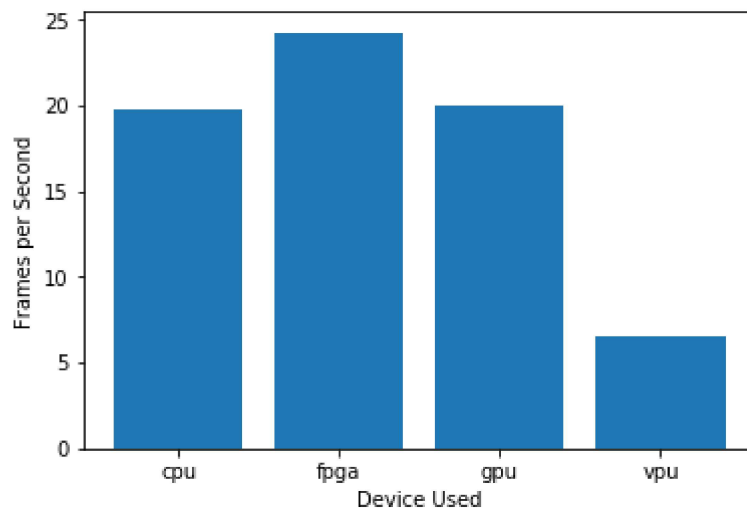
Test Results

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).





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

VPU/NCS2 is the recommended hardware. The client requires additional hardware investment as the CPU used for surveillance will not have much available computation for inference so this eliminates CPU and IGPU. However the client also have a budget of under \$300 this eliminates the FPGA from consideration as it is grossly over budget.

