

Choose the Right Hardware

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)
Field Programmable Gate Array (Intel Mustang-F100-A10)

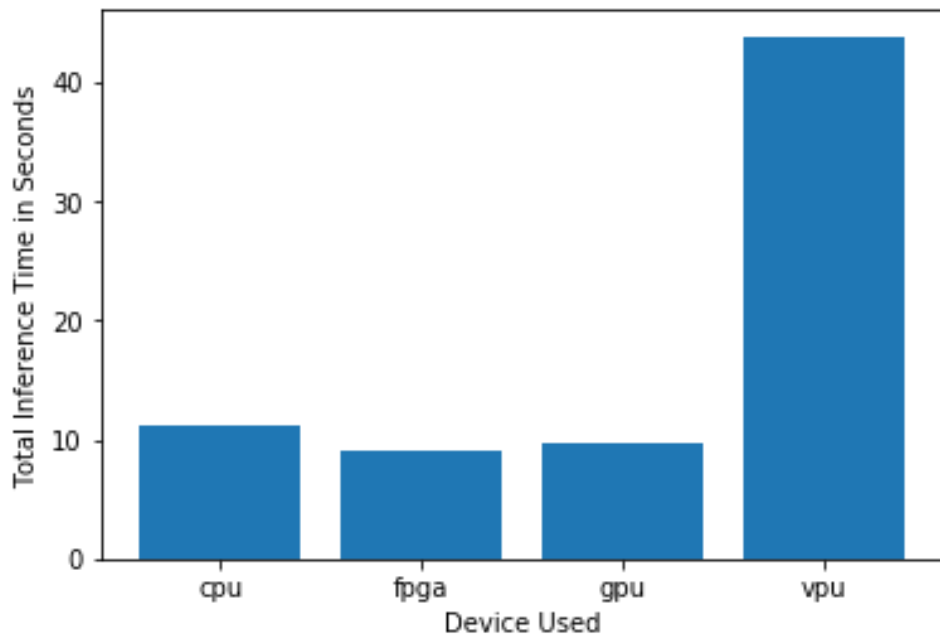
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
<i>The client wants to detect number of people in factory line, hence processing 30 to 35 frames per second, i.e. each frame needs to be processed in 0.03ms or 30 microseconds</i>	<i>Intel's Arria 10 FPGA i.e. Intel Mustang-F100-A10 can provide this processing speed</i>
<i>The client wants that the system should be flexible, so that it can be reprogrammed and optimized to quickly detect flaws in different chip designs</i>	<i>FPGA are chips designed with maximum flexibility, i.e. they can be re-programmed as and when need in the field, even after manufacturing and development</i>
<i>The client wants the system should have a life span of 5-10 years</i>	<i>Intel Mustang-F100-A10 have a guaranteed availability of 10 years, from start of production.</i>

Queue Monitoring Requirements

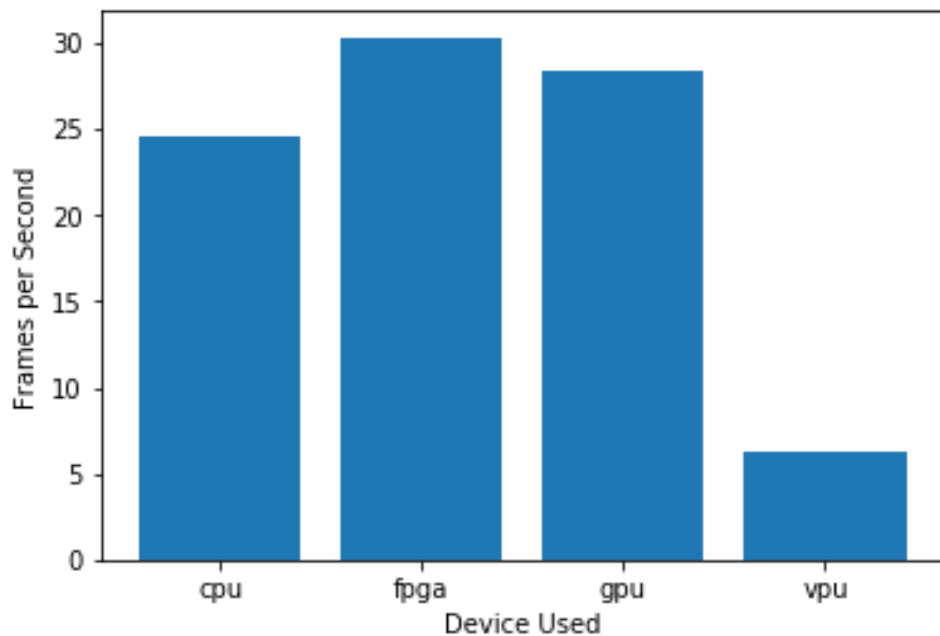
Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or Int8)	FP16

Test Results

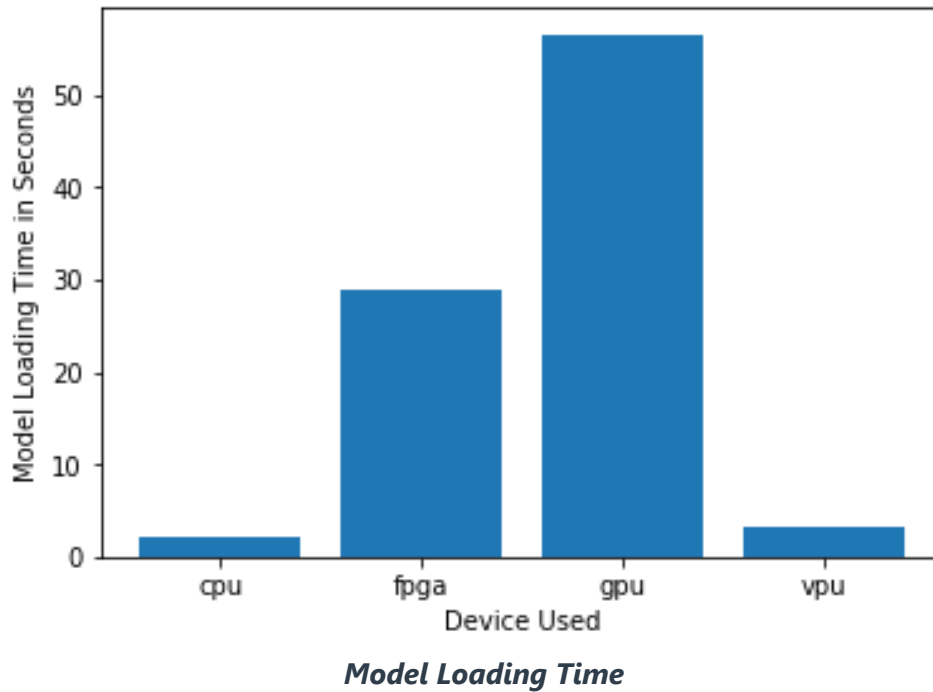
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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

So, the final hardware recommendation is FPGA (Intel Mustang F-100-A10)

Client's 1st requirement was to be able to process 30 to 35 frames/second, According to the graph, we can see that FPGA was able to process 30 frames/second.

Next requirement was that the system should be flexible enough that they can be reprogrammed and optimized to quickly detect flaws, which can be easily achieved by FPGA.

Last requirement was to have a life span of 5-10 years, which can be achieved by Intel Mustang-F100-A10, as it has a guaranteed availability of 10 years, from start of production.

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)
Central Processing Unit (Intel core i7)

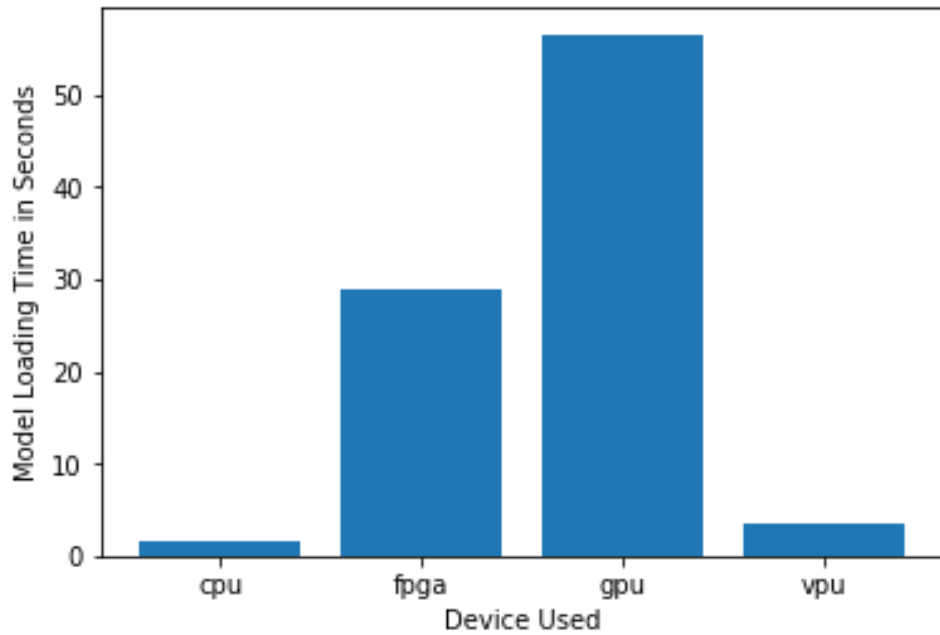
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
<i>The client does not want to spend more</i>	<i>Since the client has an Intel i7 core CPU, hence the client does not need to spend any more money</i>
<i>The client want minimum power consumption</i>	<i>Since the PC client is using has a TDP of 130 W, which is enough to handle the power consumption requirement of the system</i>
<i>The client does not require a very high processing time, i.e. each frame can take some seconds to process</i>	<i>Intel core i7 has a clock speed of around 4 GHz, i.e. each frame can be processed in less than a nano second</i>

Queue Monitoring Requirements

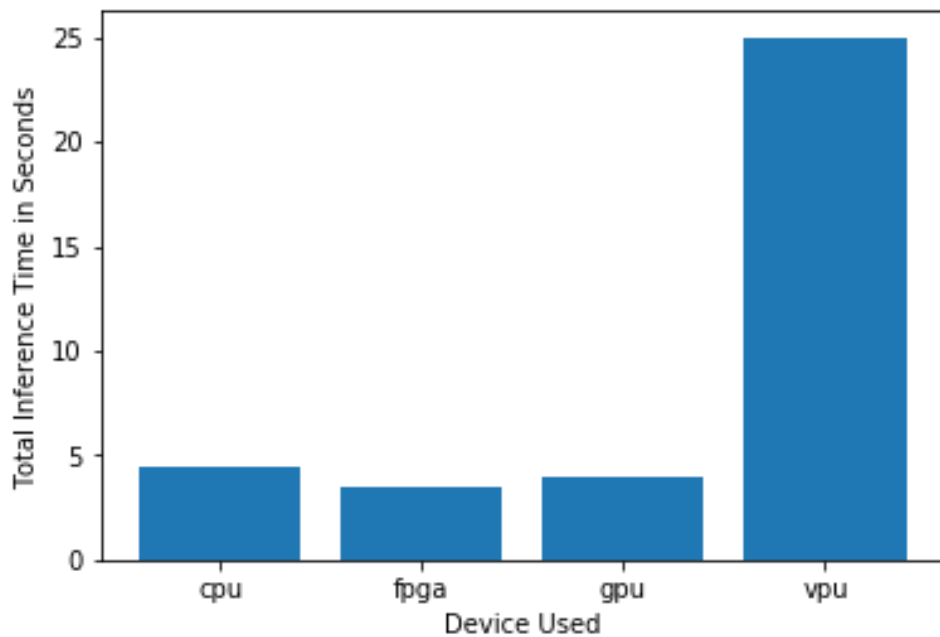
Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or Int8)	FP32

Test Results

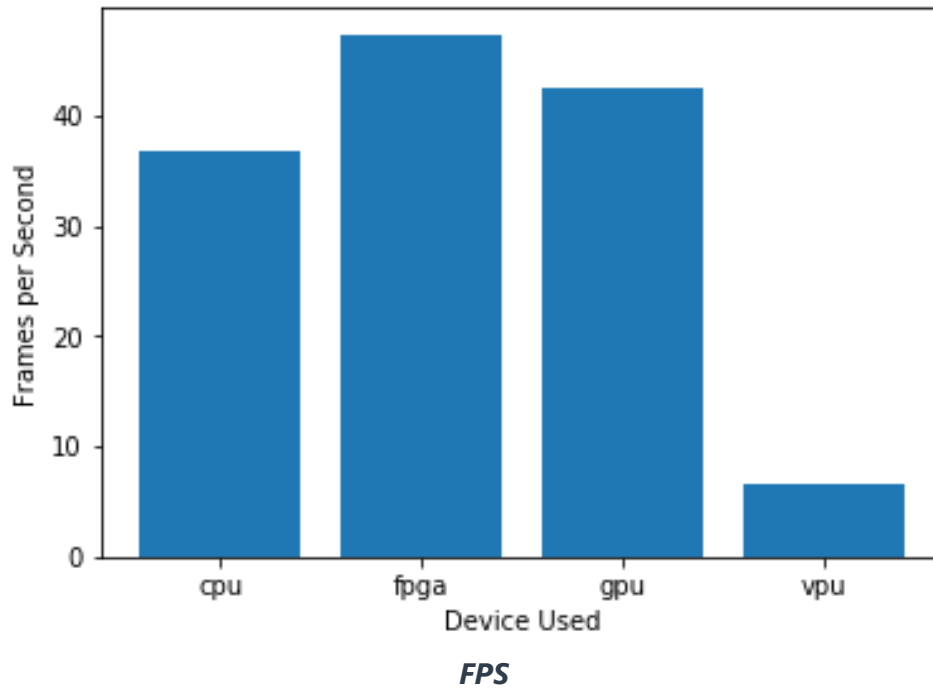
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



Model Load Time



Inference Time



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

So, the final hardware recommendation is CPU (Client's Intel Core i7)

Since the client has an Intel i7 core CPU, hence the client does not need to spend any more money

Since the PC client is using has a TDP of 130 W, which is enough to handle the power consumption requirement of the system

The FPS provide by CPU is more than 35 frames/second, which is more than enough for the client to process the frames.

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)
<i>Vision Processing Unit (Intel Neural Compute Stick 2)</i>

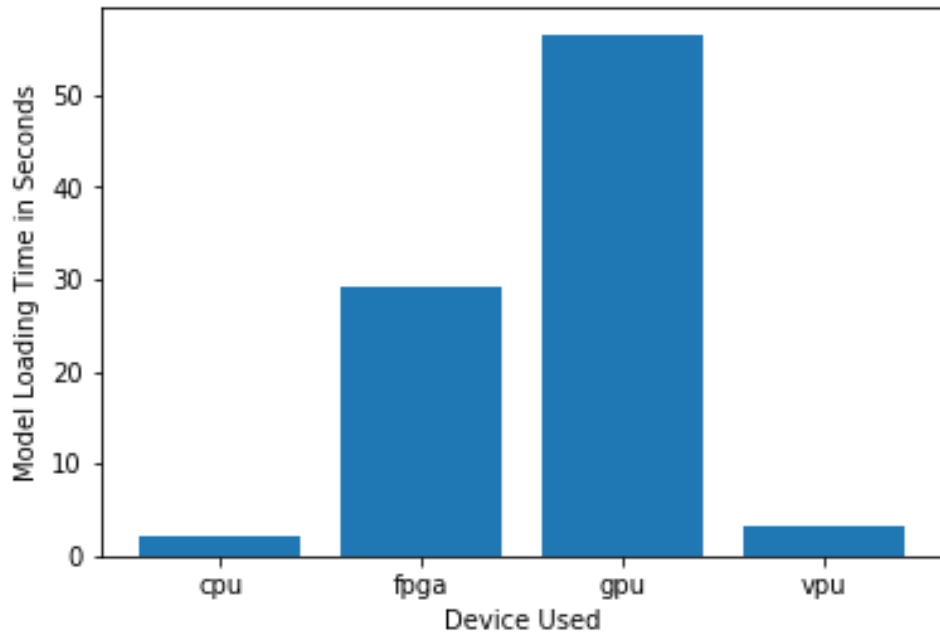
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
<i>The client want to spend at max \$300/machine</i>	<i>Price of Intel Neural Compute Stick 2 is currently around \$125</i>
<i>The client want to save on the power consumptions</i>	<i>Power consumed by Intel NCS2 is less than 1 W</i>
<i>Need high processing speed, as they have to process feed from 7 different cameras</i>	<i>Intel NCS2 has a processing speed of 4 Teraflops per second, which will be sufficient for processing 7 different camera feeds</i>

Queue Monitoring Requirements

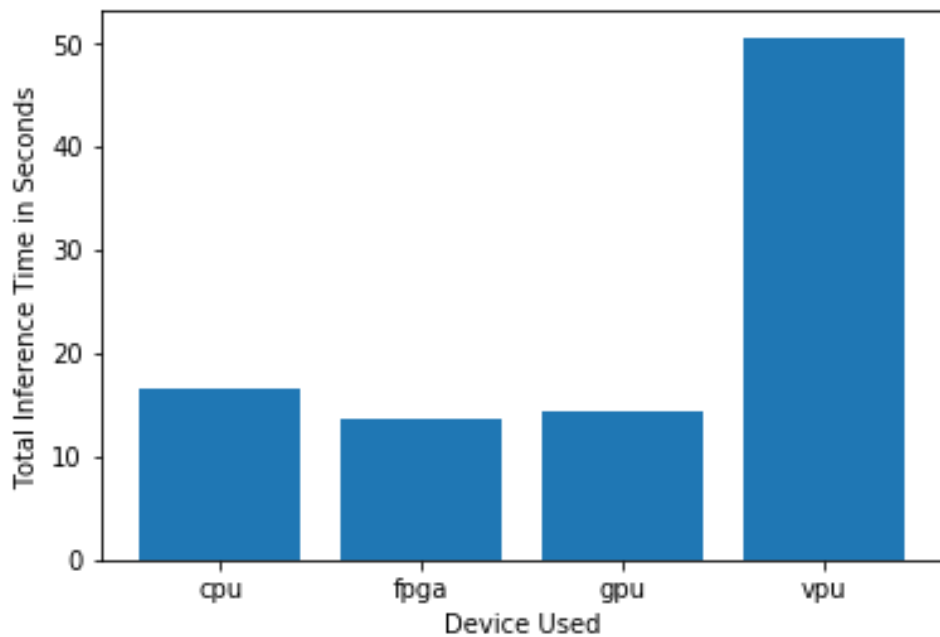
Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or Int8)	FP16

Test Results

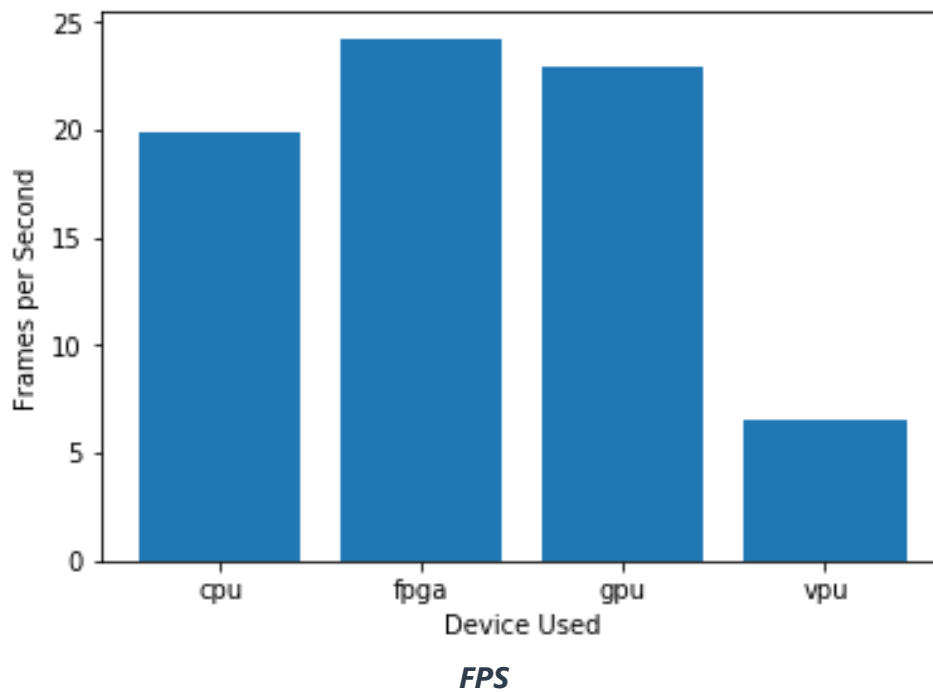
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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

The final hardware specification is VPU (Intel Neural Compute Stick 2)

Price of IGPU is currently around \$125, which is less than the client's requirement of less than \$300/machine.

Power consumed by IGPU is less than 1W which matches client's requirement, as client uses a PC, and normally a PC consumes 5-250 Watts.

Intel NCS2 was able to process at the speed of 5 FPS, Intel NCS2 has a processing speed of 4 Teraflops per second, which will be sufficient for processing 7 different camera feeds