

# 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

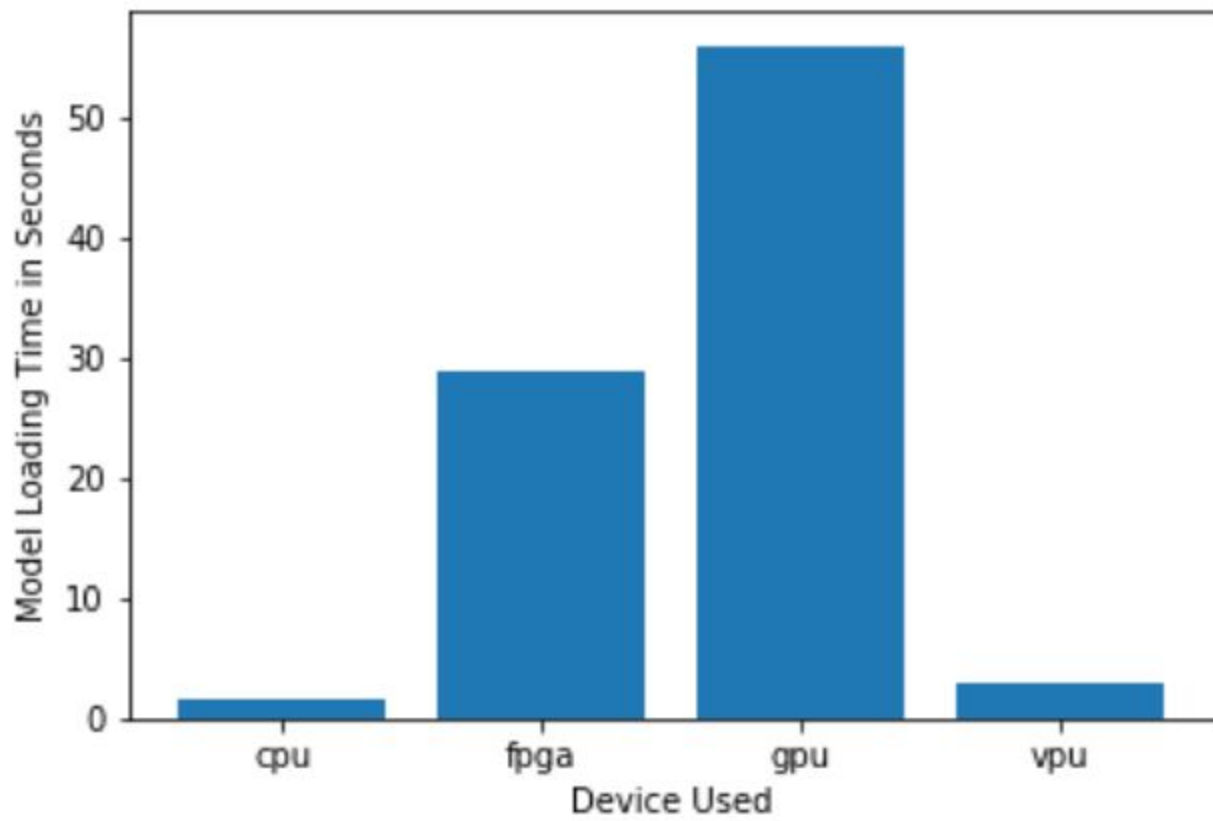
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
"Mr. Vishwas would like to be able to repurpose the system to address a second issue."	FPGAs are flexible and reprogrammable by nature.
"Naomi Semiconductors has plenty of revenue to install a quality system"	FPGAs are expensive. But the company can afford them.
"they would ideally like it to last for at least 5-10 years."	FPGAs last long enough.

#### Queue Monitoring Requirements

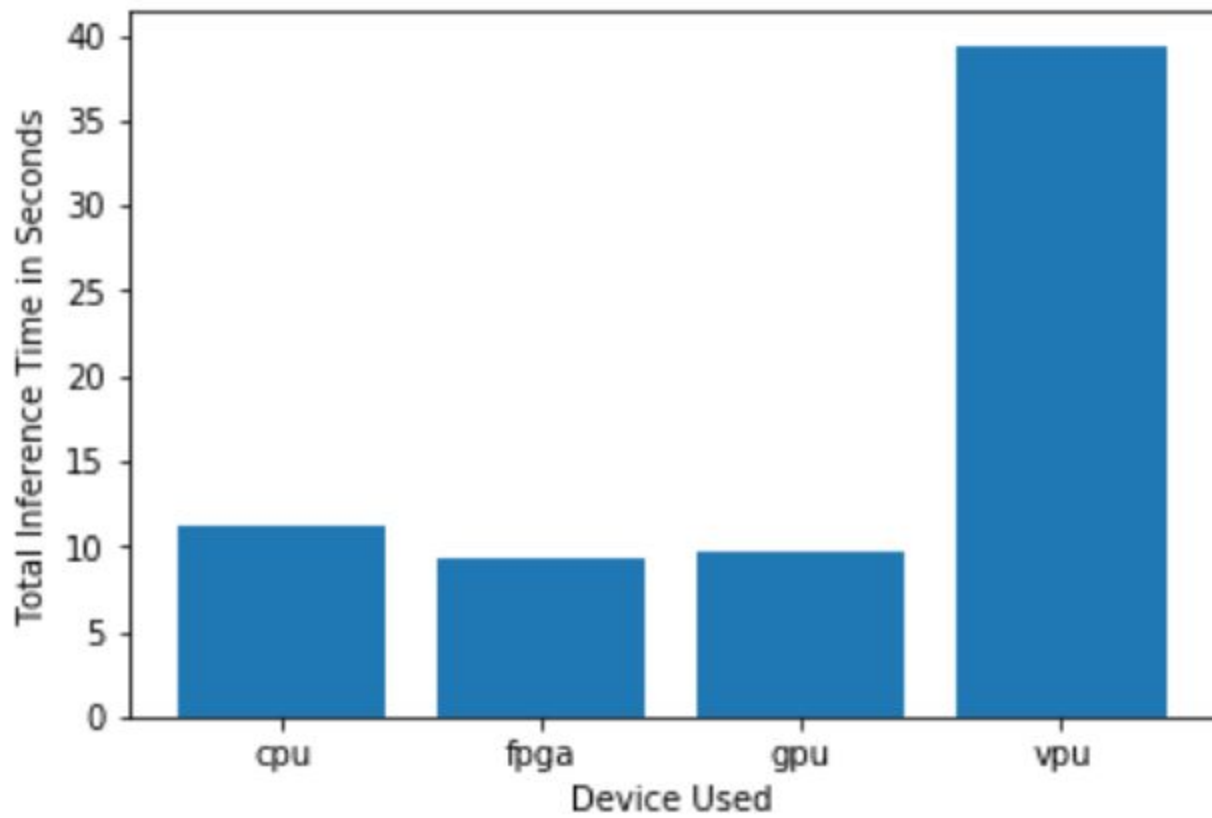
Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or Int8)	FP16

#### Test Results

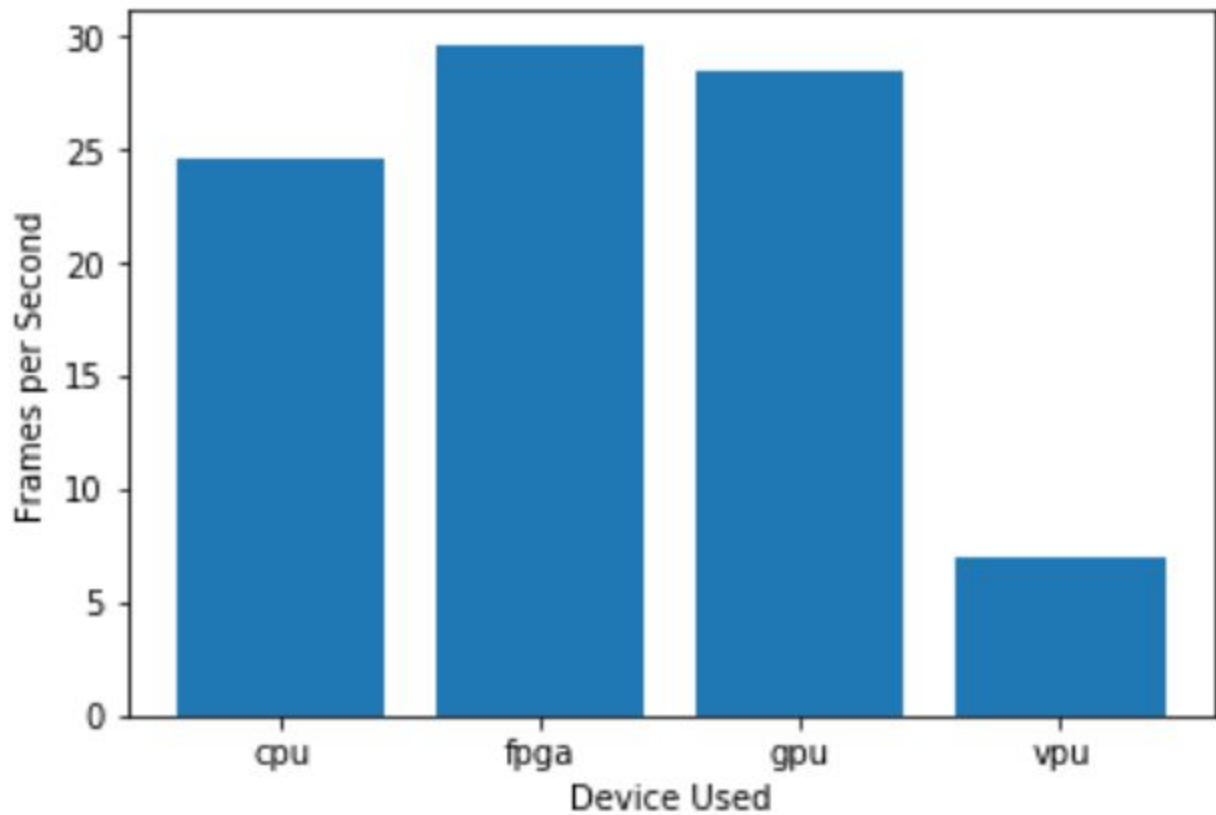
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



*Model Load Time*



### Inference Time



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

*In all 3 scenarios, the faster the hardware, the better.  
By observing the graphs on inference time and frames per second, FPGA is the fastest hardware for the task.  
Given the company has enough funding and also needs reprogrammable and durable hardware, FPGA is the right choice for Scenario 1.*

## 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

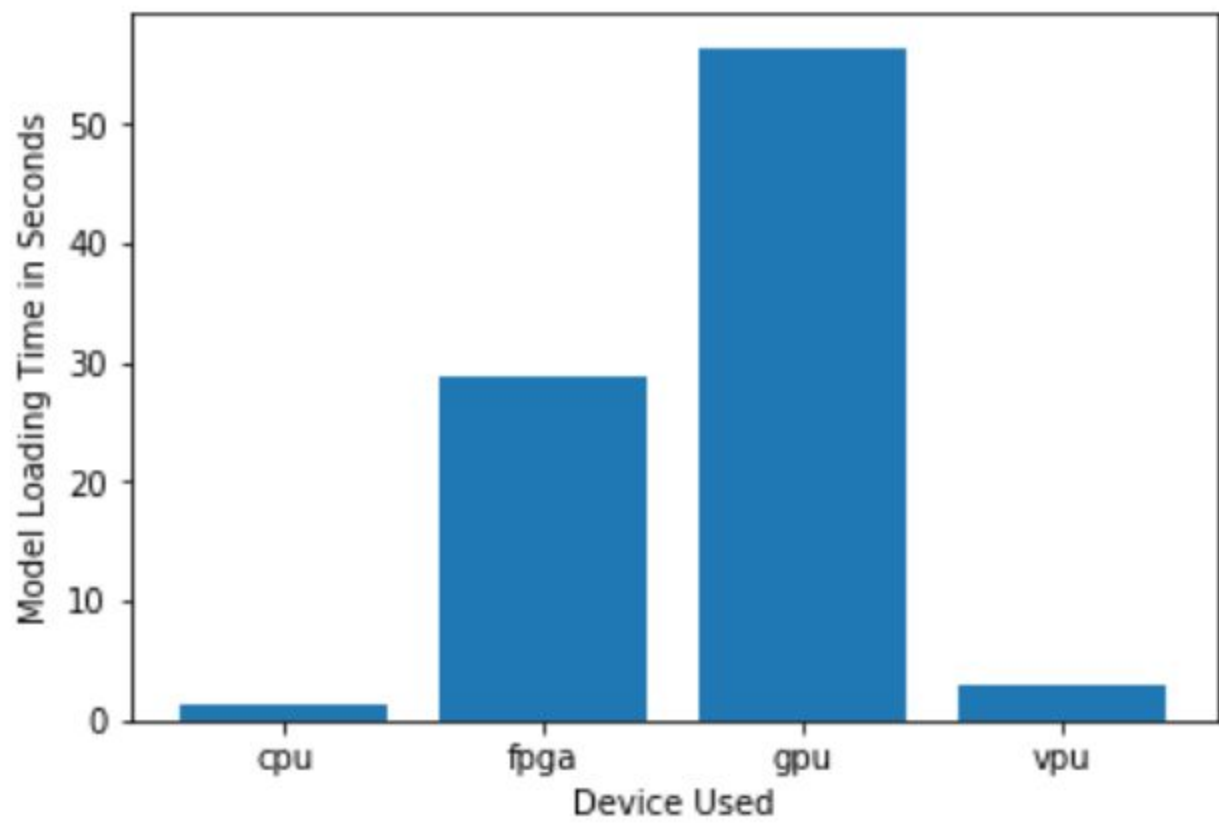
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
<i>“Most of the store's checkout counters already have a modern computer, each of which has an Intel i7 core processor. Currently these processors are only used to carry out some minimal tasks that are not computationally expensive.”</i>	<i>Mr. Lin already has 1 CPU per queue.</i>
<i>“Mr. Lin does not have much money to invest in additional hardware, and also would like to save as much as possible on his electric bill.”</i>	<i>Mr. Lin does not want to spend more money on additional hardware. So, using the preexisting CPUs is the correct choice.</i>
<i>The requirements don't specify super fast hardware.</i>	<i>CPUs have a decent performance, quite close to the best performance.</i>

## Queue Monitoring Requirements

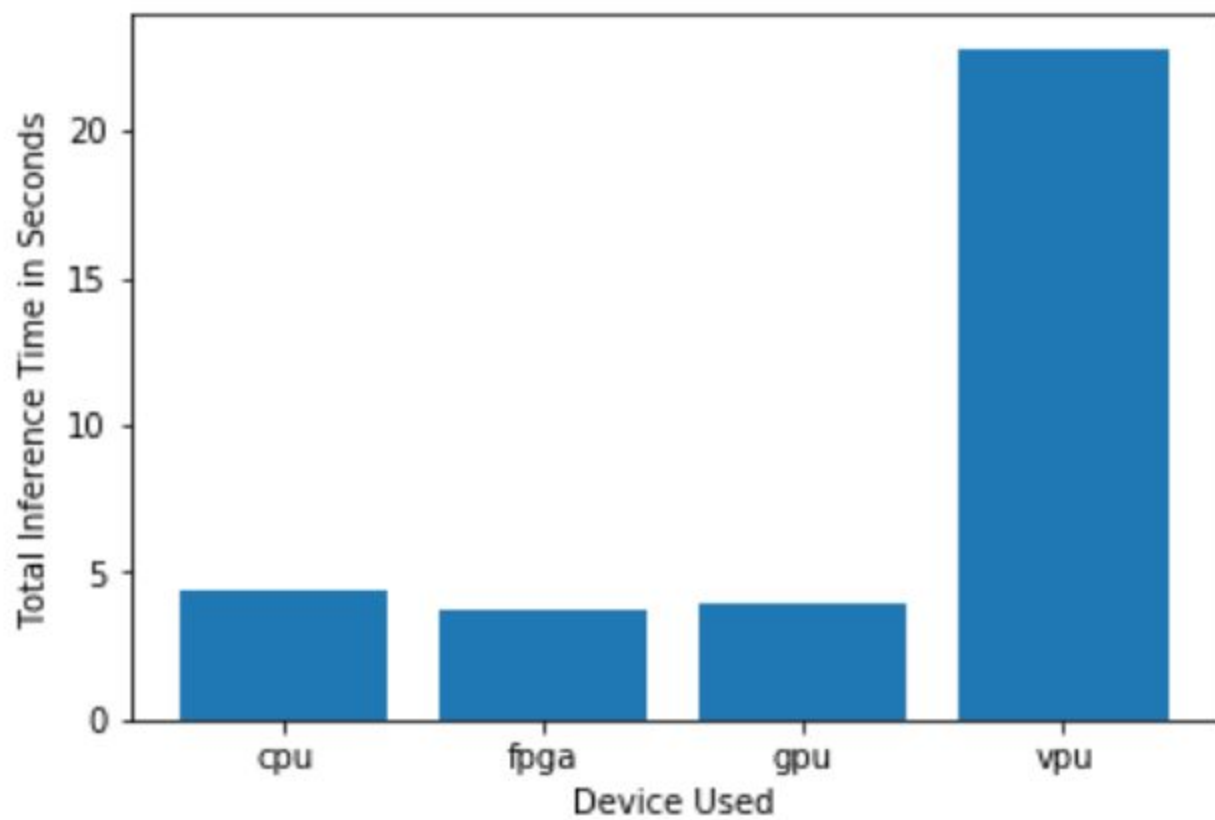
Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or Int8)	FP16

## Test Results

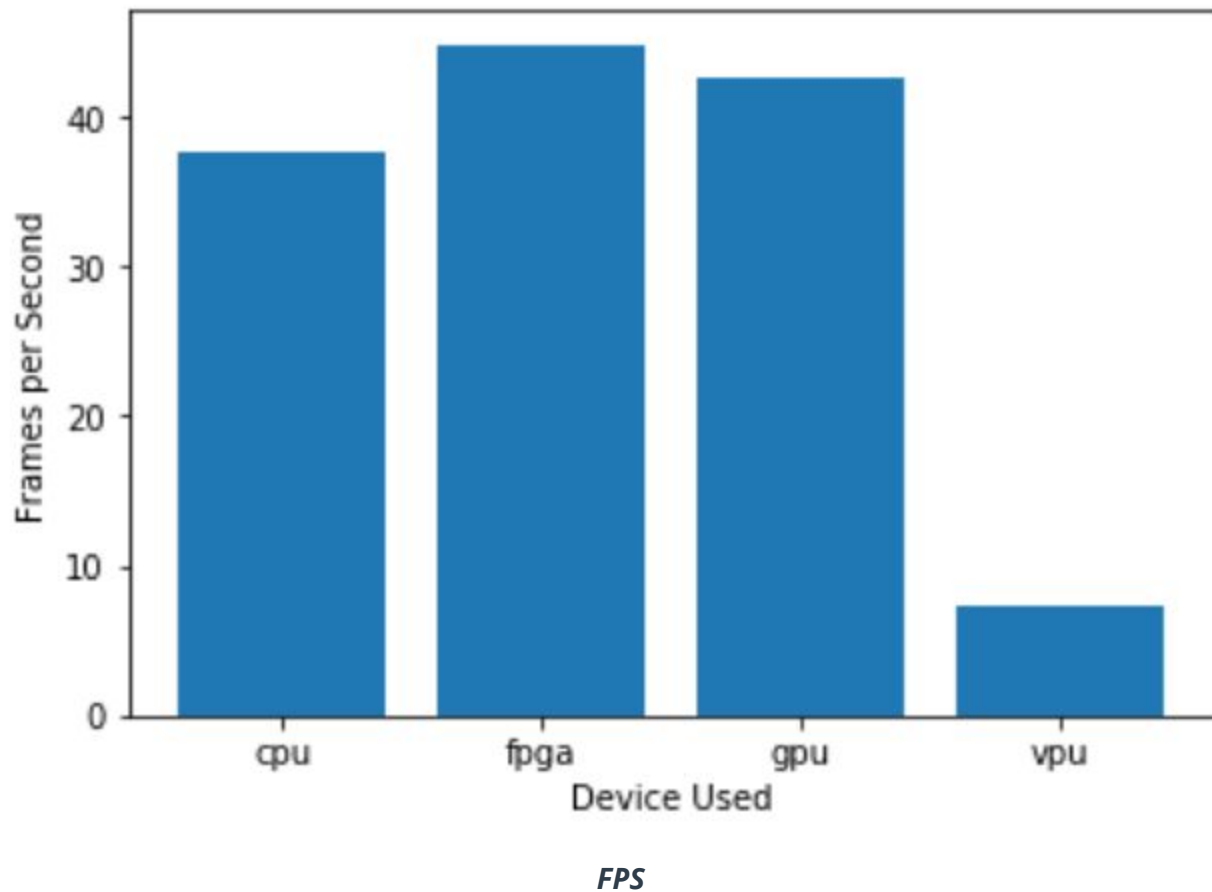
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



*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

*In all 3 scenarios, the faster the hardware, the better.*

*By observing the graphs on inference time and frames per second, CPUs have a decent performance, quite close to the best performance.*

*Mr. Lin already has 1 CPU per queue and does not want to spend more money on additional hardware. So, using the preexisting CPUs is the correct choice for Scenario 2.*

## 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)
GPU.

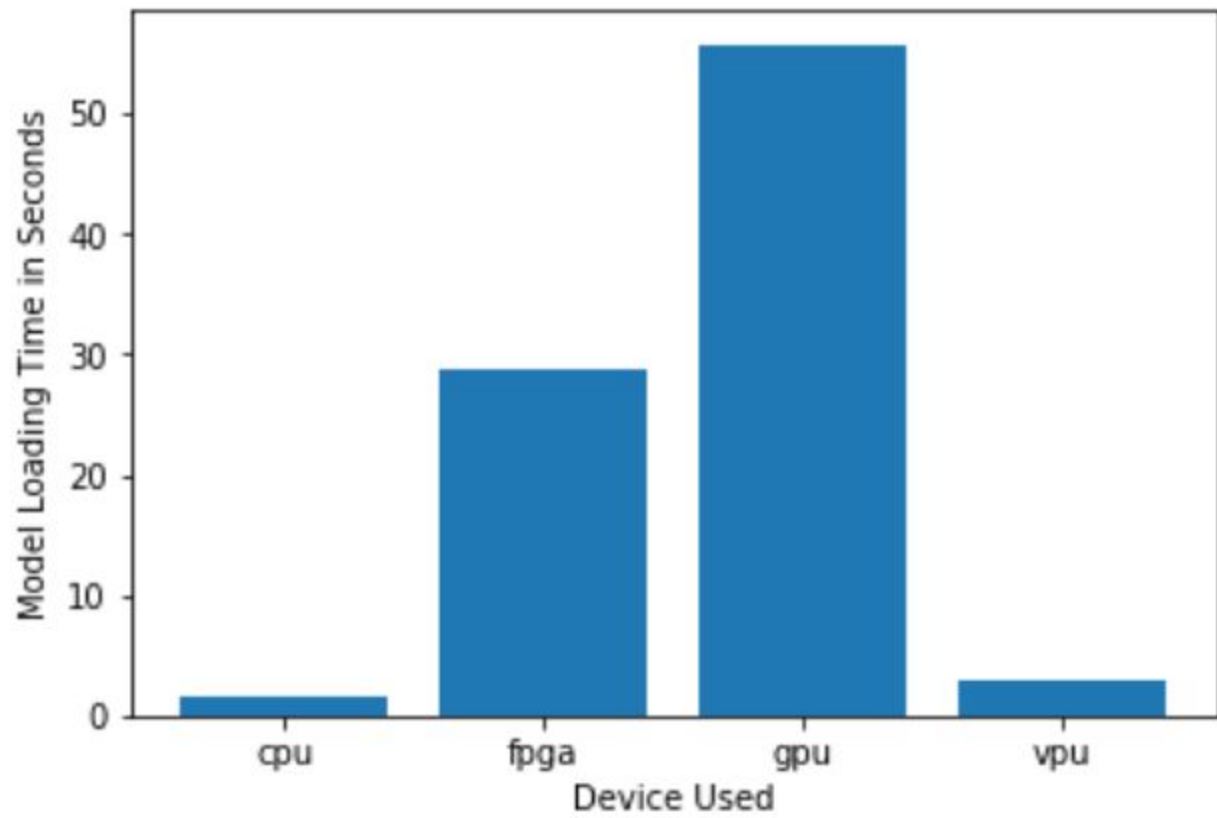
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>Example requirement:</i> The client requires a tiny device to be connected to their CPU—and their budget is only about \$100 for each device.	<i>Example explanation:</i> VPU or NCS2 is only about 27.40 mm in size and would fit in the price range.
<i>"Ms. Leah would like to automate this using an Edge AI system that would monitor the queues in real-time and quickly direct the crowd in the right manner."</i>	<i>Real-time means as fast as possible.</i>
<i>"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. 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."</i>	<i>Additional computer power is needed...</i>
<i>"Ms. Leah's budget allows for a maximum of \$300 per machine, and she would like to save as much as possible both on hardware and future power requirements."</i>	<i>...at an affordable price. Hence, the GPU. Because the VPU (NCS2) is too slow to be real-time.</i>

## Queue Monitoring Requirements

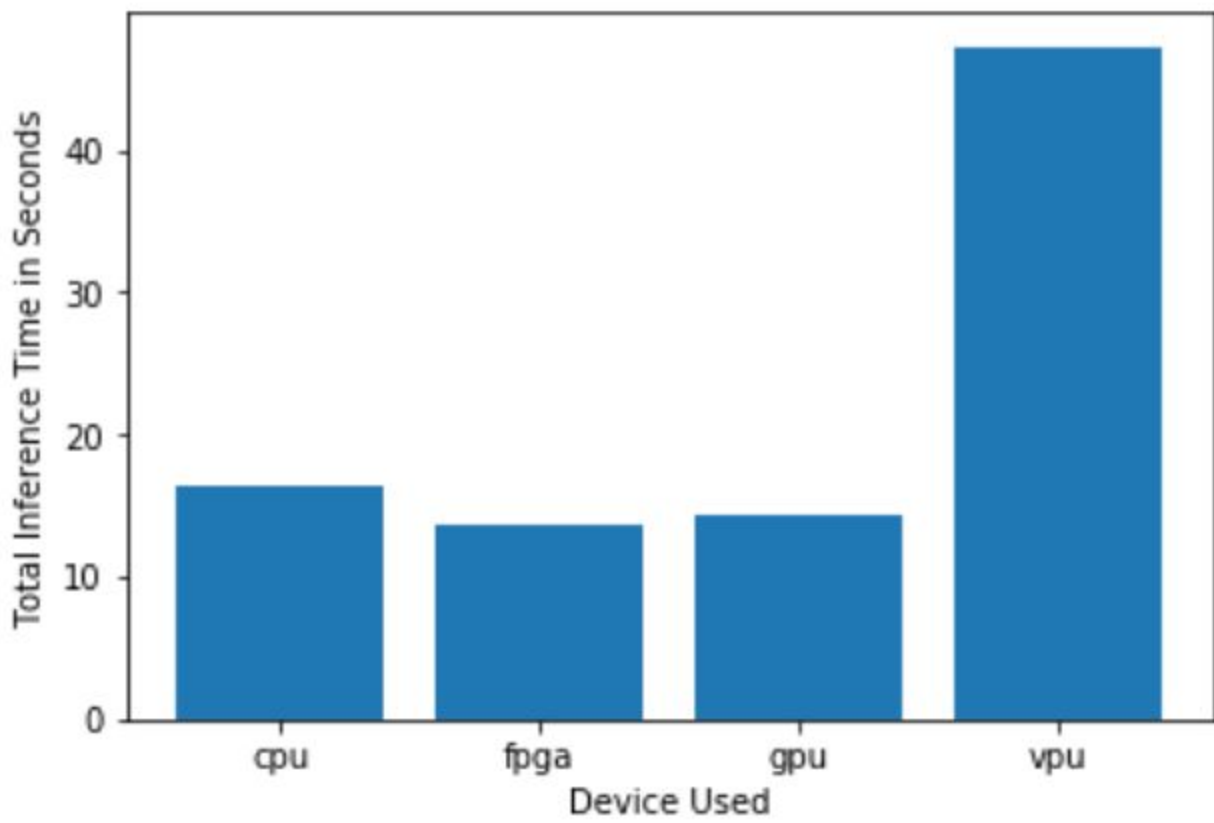
Maximum number of people in the queue	6
Model precision chosen (FP32, FP16, or Int8)	FP16

## 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).

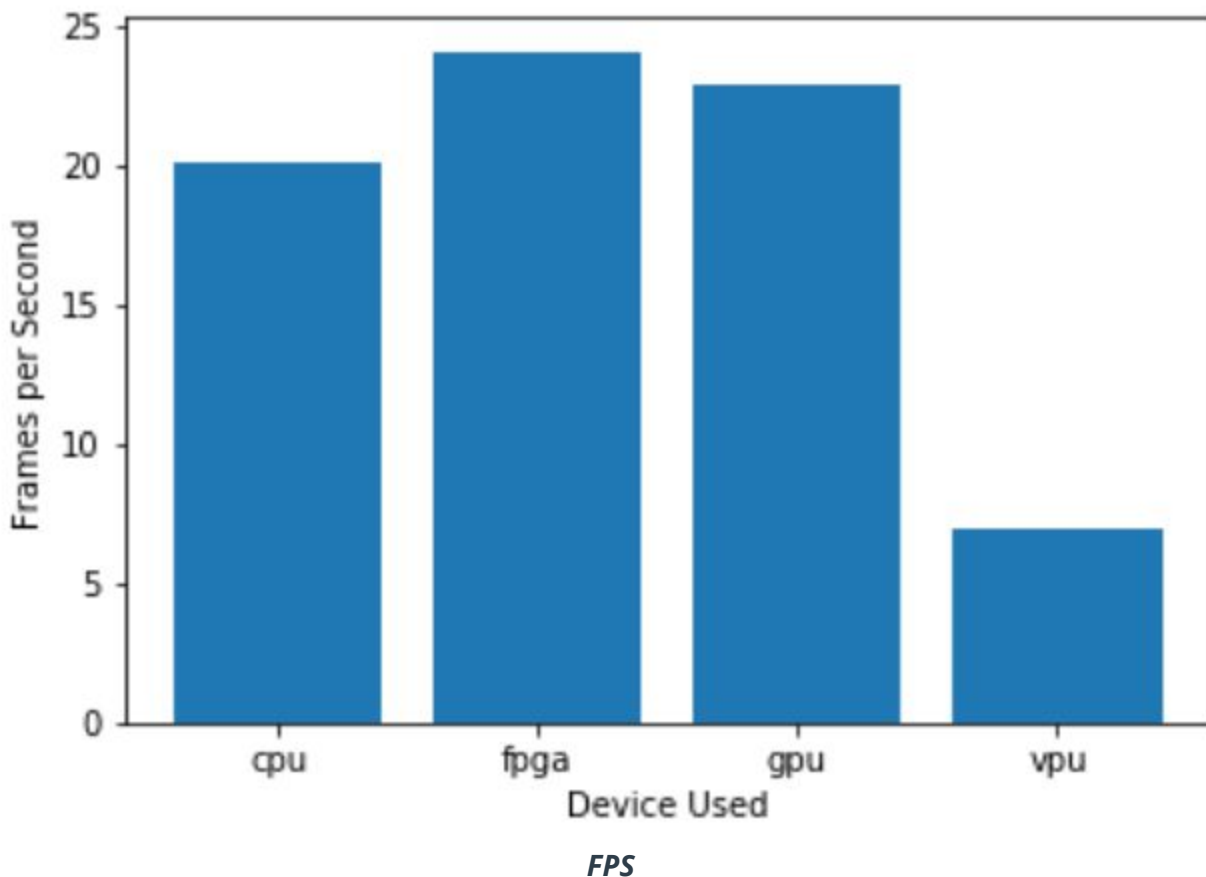


*Model Load Time*



*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 requirements state that real-time performance is needed, which means hardware should be as fast as possible.*

*Current hardware is not enough to compute real-time inferences on real-time videos. So, additional computer power is needed at an affordable price. Hence, the GPU. Because the VPU (NCS2) is too slow to be real-time.*