

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

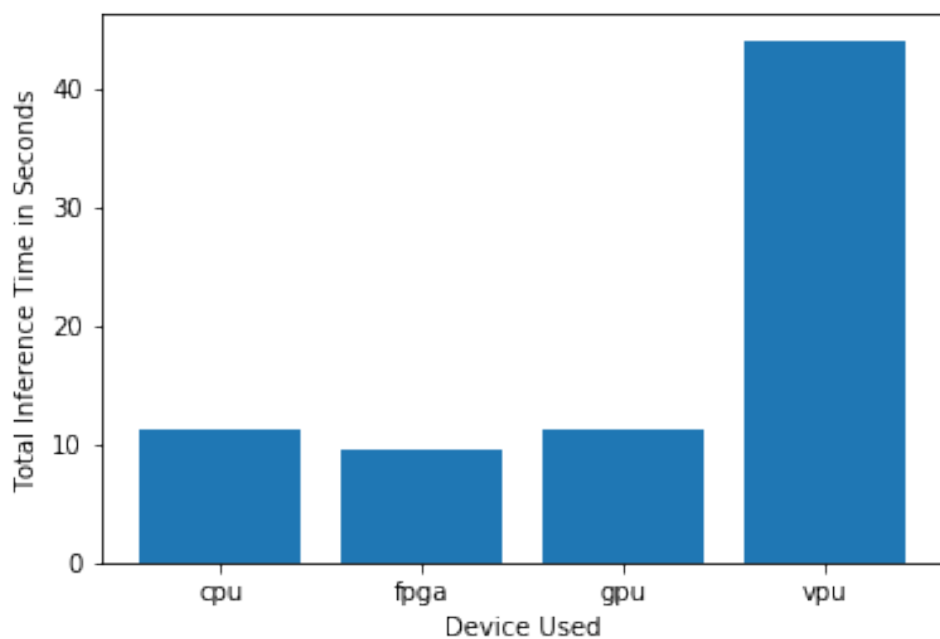
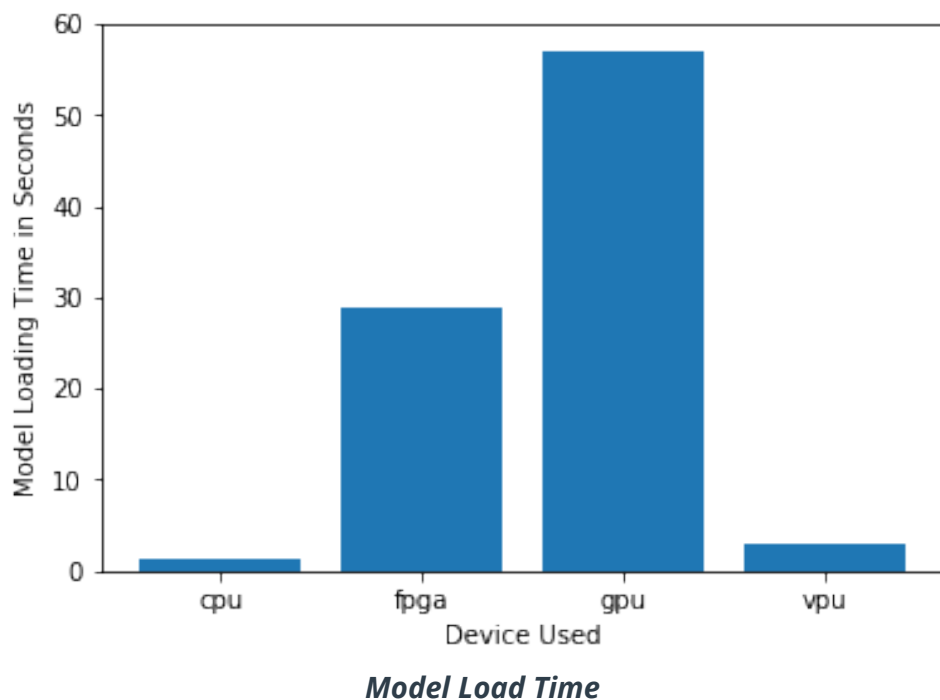
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>The client has 2 problems he seeks for solution: monitor the worker in line and after it's done the system should be able to be repurposed to detect chip flaws</i>	<i>FPGA is programmable and can be reprogrammed to fit other purposes</i>
<i>* Monitor the worker: the system should be able to run 5 times the processing work per second, corresponding inferencing $5 \times 30 \sim 35 = 150 \sim 175$ FPS * Detect chip flaw: the system should be able to run inference on video very quickly --> require fast inference</i>	<i>FPGA has in general high performance and low latency</i>
<i>Devices last at least 5-10 years</i>	<i>FPGA has in general a very long lifespan (~10 year from construction)</i>
<i>The client has potential budget to invest in the quality of system</i>	<i>It's possible to pay for FPGA</i>

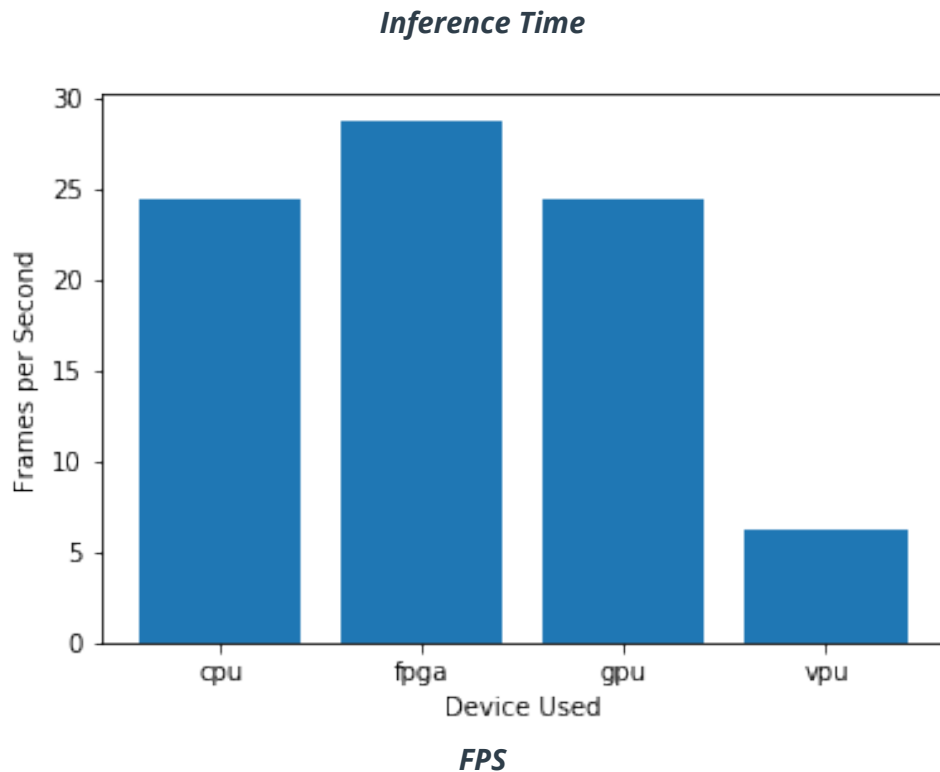
Queue Monitoring Requirements

Maximum number of people in the queue	5 (approximation in given input video)
---------------------------------------	--

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).





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 recommendation is FPGA (which is actually the set of Intel Core i5-6500te CPU and IEI Mustang-F100-A100 FPGA in testing)

After testing with cpu/ncs-2/gpu, it is to see in the graphs above that fpga has a very convincing result. It has a short inference time and can process a high FPS which meets the client's requirements of performance and speed. The client also concerns about the duration of the hardware and FPGA has in general a very long lifespan of about ~10 years. Another important requirement is also that the system should be easily reprogrammed to detect chip flaw later on. This can be done as a default feature of FPGA.

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)
<i>CPU at checkout counter</i>

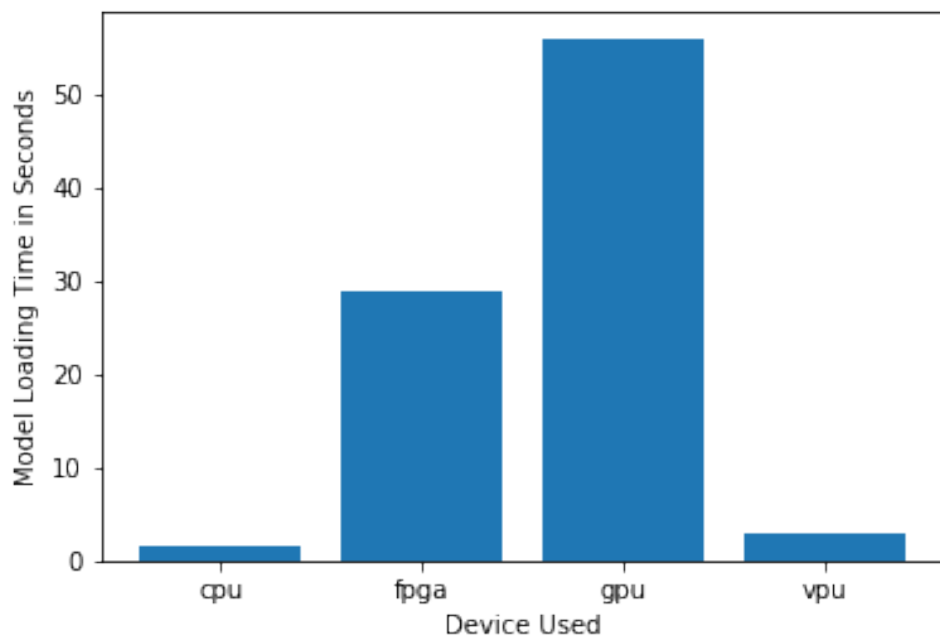
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>The client has only small budget and want to save on electric</i>	Take advantage of CPU with intel i7 at the checkout counter. It costs nothing extra and consumes not too much power

Queue Monitoring Requirements

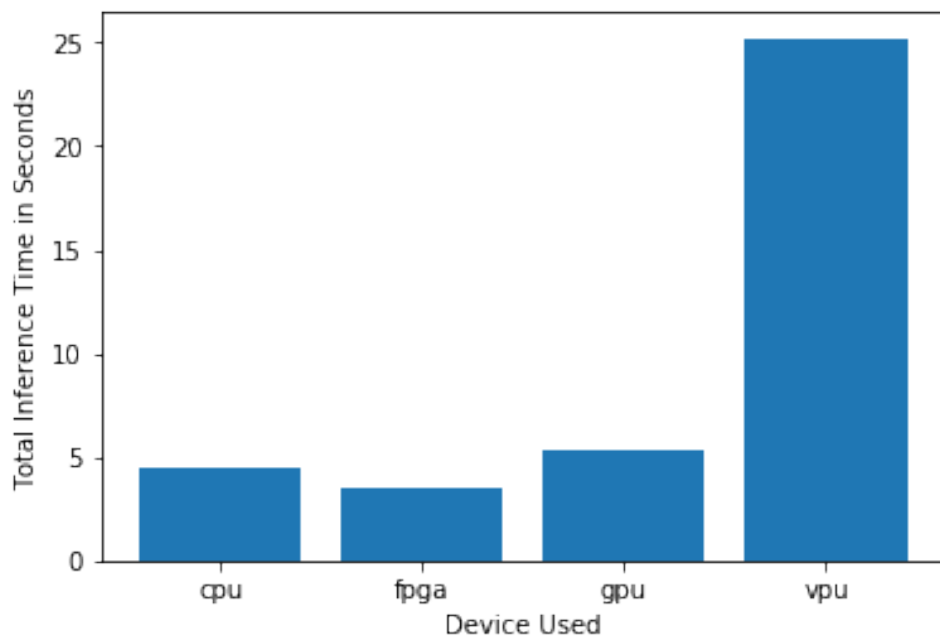
Maximum number of people in the queue	2
Model precision chosen (FP32, FP16, or Int8)	FP32

Test Results

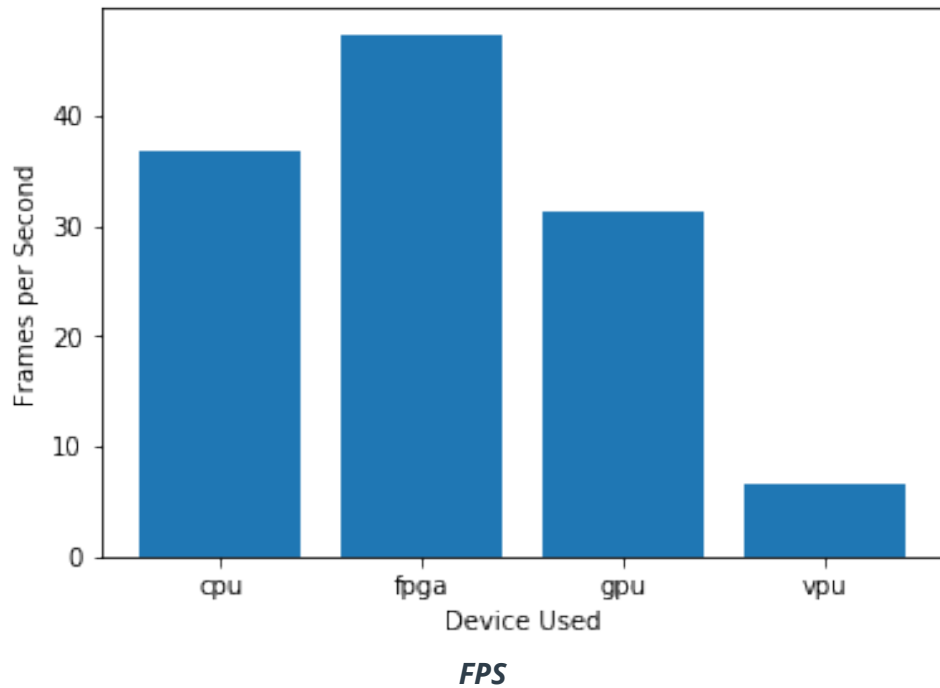
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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

The final hardware recommendation is the available CPU at checkout counter.

CPU has lowest model loading time, a small inference time and a high fps in comparison to gpu/fpga/ncs-2. Since there are available modern computers with intel i7 core processor and they are currently only used for some low-computational tasks, it's the best way to take advantage of them. So that Mr. Lin does not have to invest any money in buying new hardware and the performance in CPU is considered good enough after testing in intel-devcloud

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 (like NCS-2)

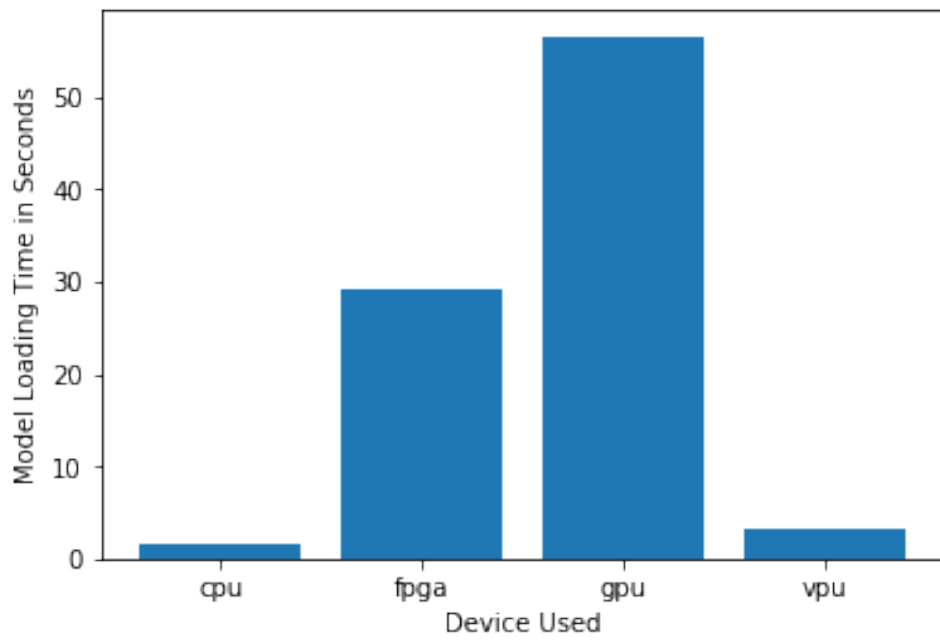
Requirement Observed (Include at least two.)	How does the chosen hardware meet this requirement?
<i>The client has a small budget and wants to save as much as possible both on hardware and future power consumption</i>	<i>VPU like NCS-2 costs a little in comparison to other hardware like FPGA or GPU and has a very low power consumption</i>

Queue Monitoring Requirements

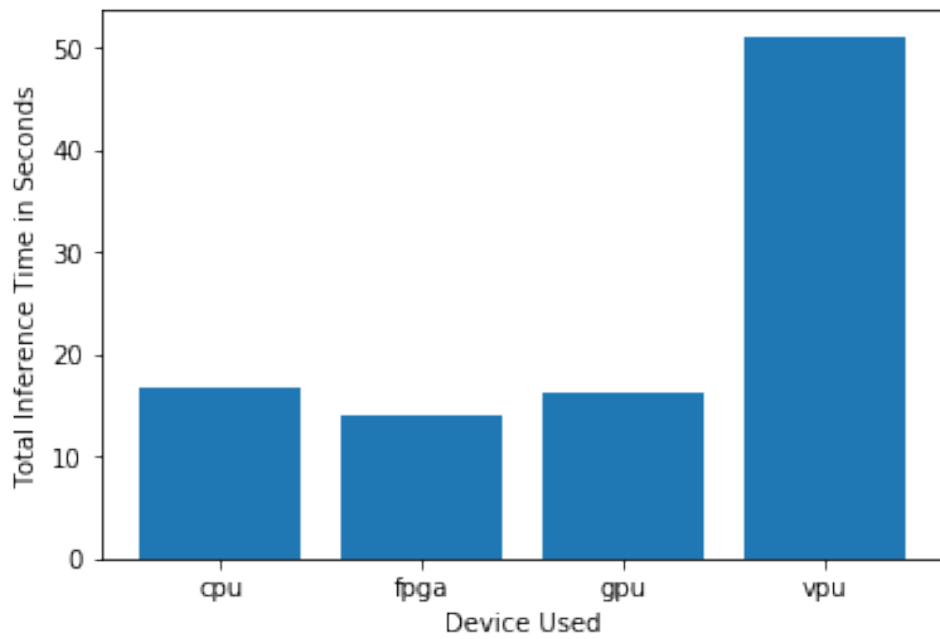
Maximum number of people in the queue	10
Model precision chosen (FP32, FP16, or Int8)	FP32

Test Results

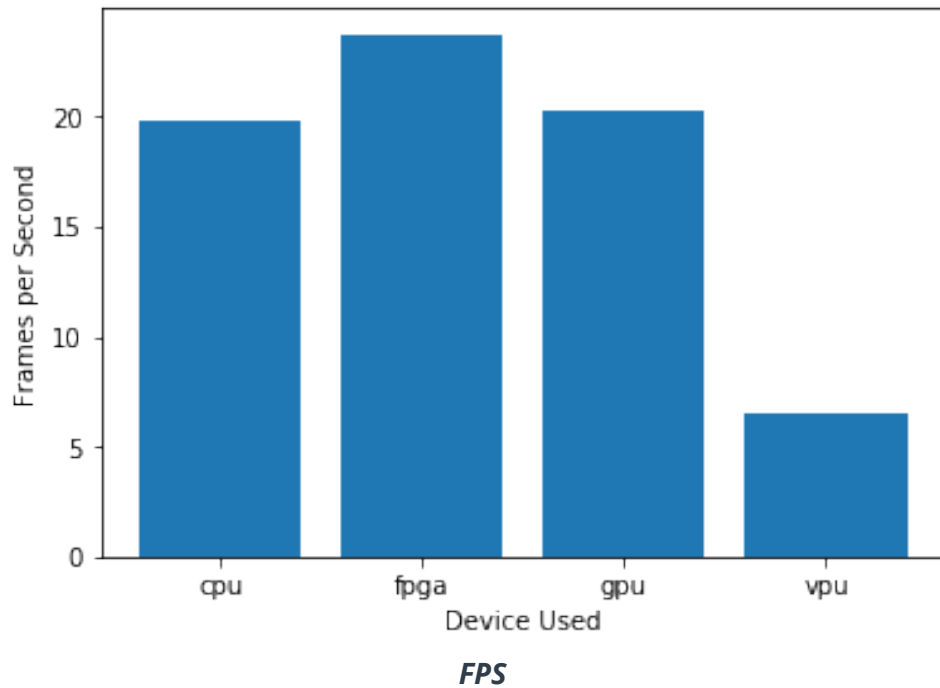
After you've tested your application on all four hardware types (CPU, IGPU, VPU, and FPGA), copy the matplotlib output showing the comparison into the spaces below. You should have three graphs (for model load time, inference time, and FPS).



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

The final hardware recommendation is VPU (NCS-2)

There are available CPUs for security purpose but it has unfortunately no significant additional processing power for inference which means new hardware need to be bought. Ms Leah budget is maximum 300\$ for a machine. Therefore it's not possible to buy FPGA even though it performs best. For this budget it's likely to invest in CPU or VPU. The performance of CPU is through testing better than VPU (faster inference and higher FPS). But since the people around the train door do not move much during the waiting time and therefore high speed is not a priority. In comparison of price among CPU and VPU, it's best to choose VPU, which fit the available budget and perform good enough for the case.