Heap buffer overflow in `QuantizedReshape`

Low **mihaimaruseac** published **GHSA-2gfx-95x2-5v3x** on May 12, 2021

Package

tensorflow, tensorflow-cpu, tensorflow-gpu (pip)

Affected versions	Patched versions
< 2.5.0	2.1.4, 2.2.3, 2.3.3, 2.4.2

Description

Impact

An attacker can cause a heap buffer overflow in `QuantizedReshape` by passing in invalid thresholds for the quantization:

```
import tensorflow as tf

tensor = tf.constant([], dtype=tf.qint32)
shape = tf.constant([], dtype=tf.int32)
input_min = tf.constant([], dtype=tf.float32)
input_max = tf.constant([], dtype=tf.float32)

tf.raw_ops.QuantizedReshape(tensor=tensor, shape=shape, input_min=input_min, input_max=input_max)
```

This is because the implementation assumes that the 2 arguments are always valid scalars and tries to access the numeric value directly:

```
const auto& input_min_float_tensor = ctx->input(2);
...
const float input_min_float = input_min_float_tensor.flat<float>()(0);
const auto& input_max_float_tensor = ctx->input(3);
...
const float input_max_float = input_max_float_tensor.flat<float>()(0);
```

However, if any of these tensors is empty, then `.flat<T>()` is an empty buffer and accessing the element at position 0 results in overflow.

Patches

We have patched the issue in GitHub commit a324ac84e573fba362a5e53d4e74d5de6729933e.

The fix will be included in TensorFlow 2.5.0. We will also cherrypick this commit on TensorFlow 2.4.2, TensorFlow 2.3.3, TensorFlow 2.2.3 and TensorFlow 2.1.4, as these are also affected and still in supported range.

For more information

Please consult our security guide for more information regarding the security model and how to contact us with issues and questions.

Attribution

This vulnerability has been reported by Ying Wang and Yakun Zhang of Baidu X-Team.

Severity

Low

CVE ID

CVE-2021-29536

Weaknesses

No CWEs