
APACHE
nifi
COMMUNITY
OVER CODE
APACHECON Europe, Oct 23rd, 2019
Running visual quality inspection at the edge with Apache NiFi & MiNiFi
Pierre Villard - @pvillard31
Google Cloud

Pierre Villard // @pvillard31

Customer Engineer @ Google Cloud
Committer and PMC member for Apache NiFi (in the community since 2015)

Twitter/Github – @pvillard31
Blog – https://www.pierrevillard.com/

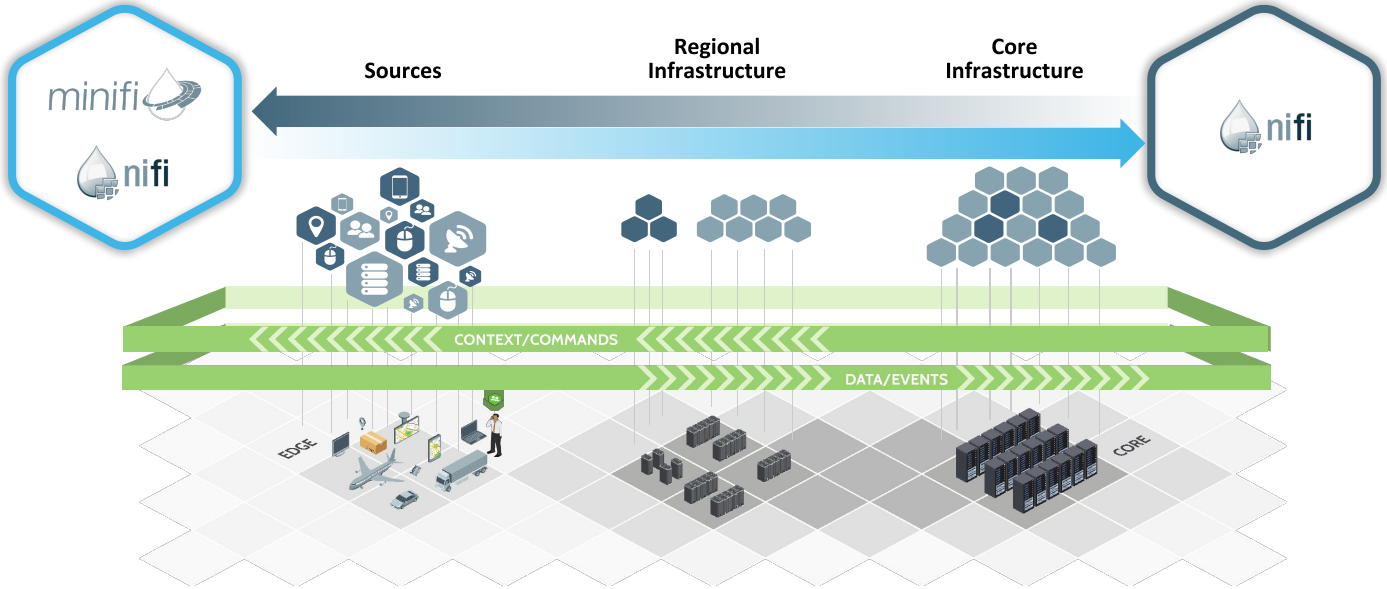
NiFi - a software developed 13y ago by the NSA

2006
NiagaraFiles (NiFi) was first incepted at the National Security Agency (NSA)

November 2014
NiFi is donated to the Apache Software Foundation (ASF) through NSA's Technology Transfer Program and enters ASF's incubator.

July 2015
NiFi reaches ASF top-level project status

What is NiFi used for?



The NiFi ecosystem

- NiFi - Powerful and scalable directed graphs of data routing, transformation, and system mediation logic.
- MiNiFi (Java version) - Complementary data collection approach that supplements the core tenets of NiFi in dataflow management, focusing on the collection of data at the source of its creation.
- MiNiFi (C++ version) - The C++ implementation is an additional implementation to the one in Java with the aim of an even smaller resource footprint. Perspectives of the role of MiNiFi should be from the perspective of the agent acting immediately at, or directly adjacent to, source sensors, systems, or servers.
- NiFi Registry - Complementary application that provides a central location for storage and management of shared resources across one or more instances of NiFi and/or MiNiFi.
- NiFi C2 Server - Command and control server to manage many disparate agents running on all sorts of devices, to coordinate their work and to push out revised flows/configurations.
- NiFi Fluid Design System - Atomic reusable platform providing consistent set of UI/UX components.

300+ processors for deeper ecosystem integration

- FTP
- SFTP
- HL7
- UDP
- XML
- ...
- HTTP
- WebSocket
- Email
- HTML
- Image
- Syslog
- AMQP

Hash	Encrypt	GeoEnrich
Merge	Tail	Scan
Extract	Evaluate	Replace
Duplicate	Execute	Translate
Split	Fetch	Convert
...	...	
Route Text	Distribute Load	
Route Content	Generate Table Fetch	
Route Context	Jolt Transform JSON	
Control Rate	Prioritized Delivery	

The Apache way: community over code

Subscribe to the mailing lists:	https://nifi.apache.org/mailing_lists.html users@nifi.apache.org & dev@nifi.apache.org
Open and comment JIRAs:	https://issues.apache.org/jira/projects/NIFI
Contribute code:	https://nifi.apache.org/developer-guide.html https://cwiki.apache.org/confluence/display/NIFI/Contributor+Guide https://issues.apache.org/jira/projects/NIFI/issues/
Get involved in the code review process:	https://github.com/apache/nifi https://github.com/apache/nifi-registry https://github.com/apache/nifi-minifi https://github.com/apache/nifi-minifi-cpp https://github.com/apache/nifi-fds

The Apache NiFi community in few numbers

535+ members on the Slack channel

260+ contributors on Github across the repositories

45 committers in the Apache NiFi community

Apache NiFi 1.10.0 to be released soon (RC vote in progress!)

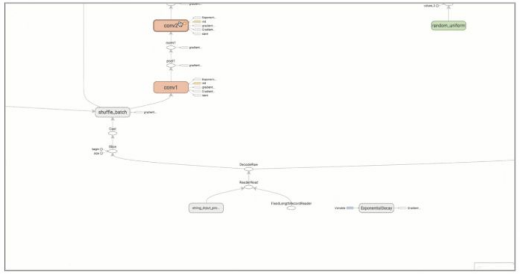
1M+ docker pulls of the Apache NiFi image

Visual quality inspection: Detect broken cookies

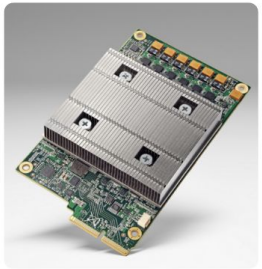
The ML spectrum in GCP

AutoML

TensorFlow



CloudML



Scale, No-ops Infrastructure

TPU - 7 years ahead of GPU in terms of price/performance

Google Trained Models

Cloud Translate

Cloud Vision

Cloud Natural Language

Cloud Speech

Use/extend OSS SDK

Build custom models

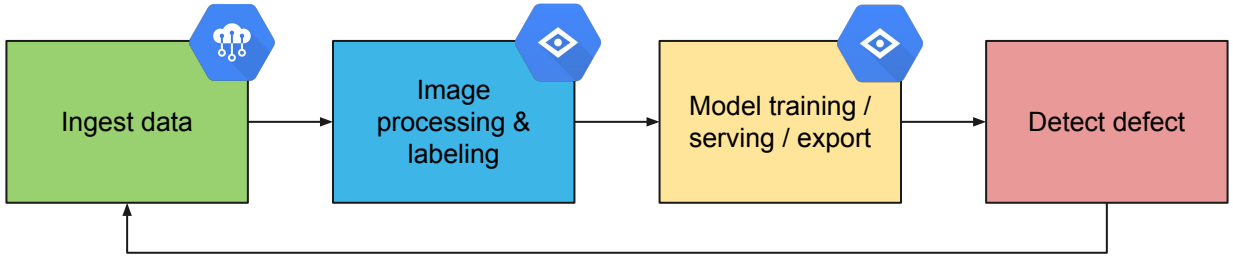
Use pre-built models

ML researcher

Data Scientist

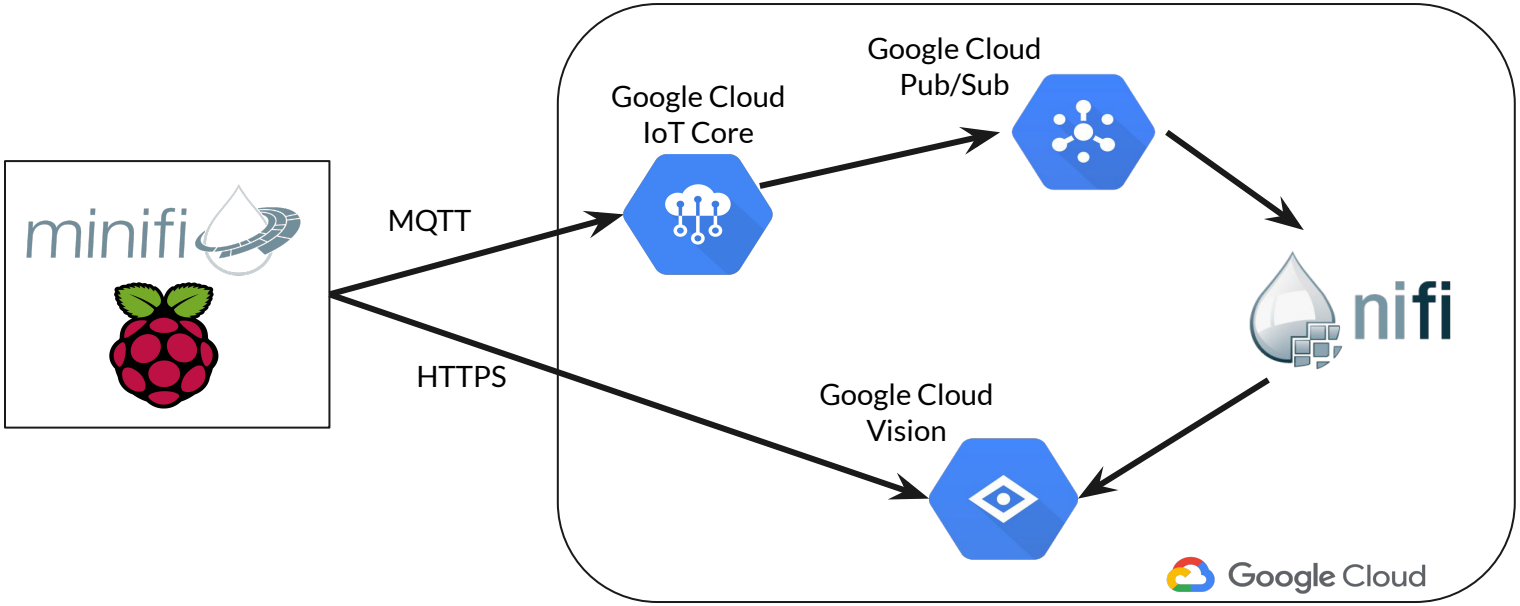
App Developer

Continuous model retrain

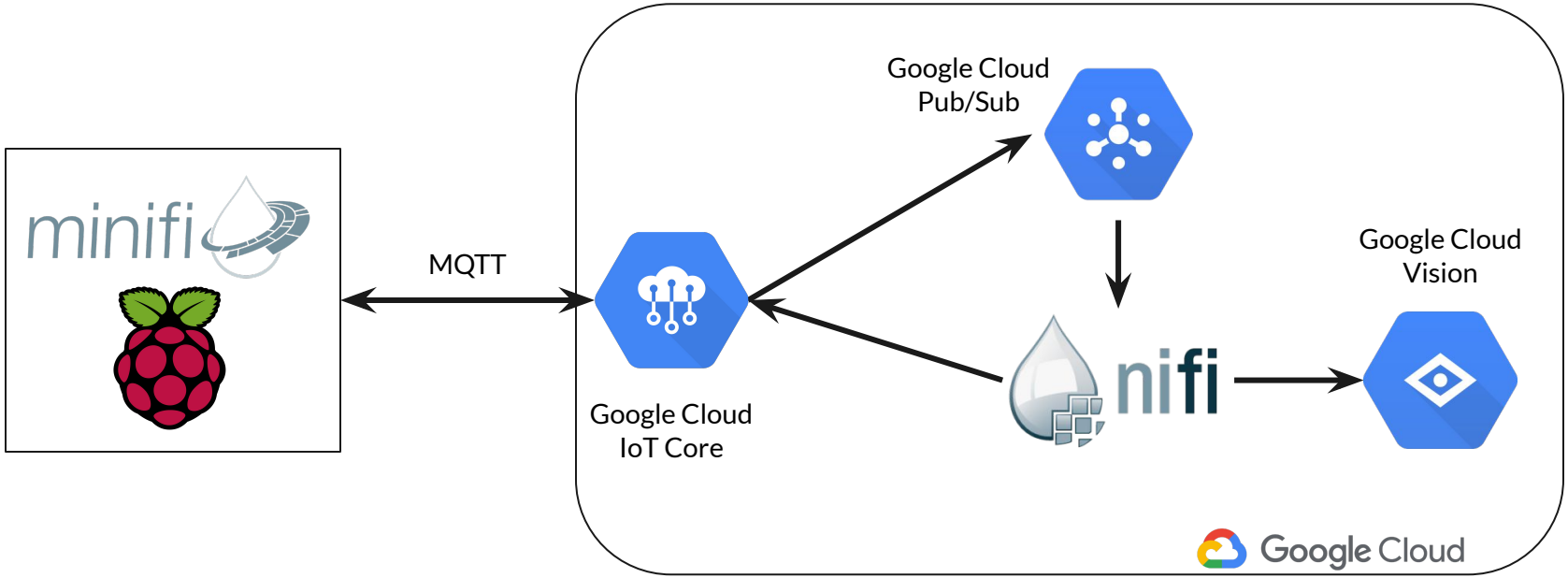


- Automatically train customized ML models in the cloud
- Efficiently acquire images, label images, deploy model and run inference
- Continuously refresh models using fresh data from the production lines

Architecture #1 - training & inference in the cloud

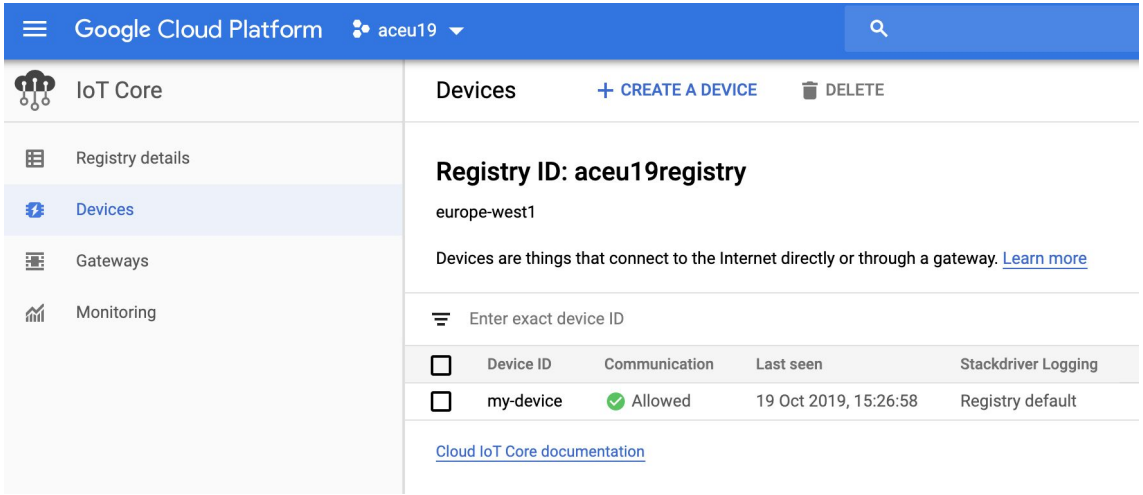


Architecture #2 - training in the cloud & inference at the edge

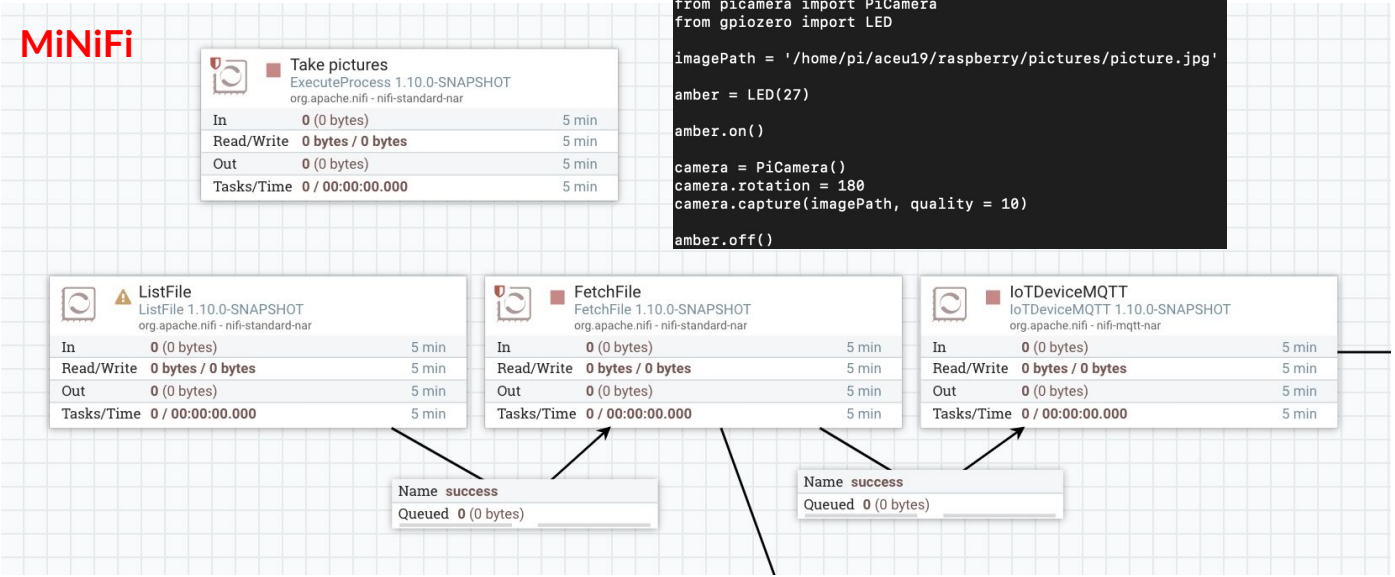


Collect & label data to initialize a dataset

Register my device in Google Cloud IoT Core

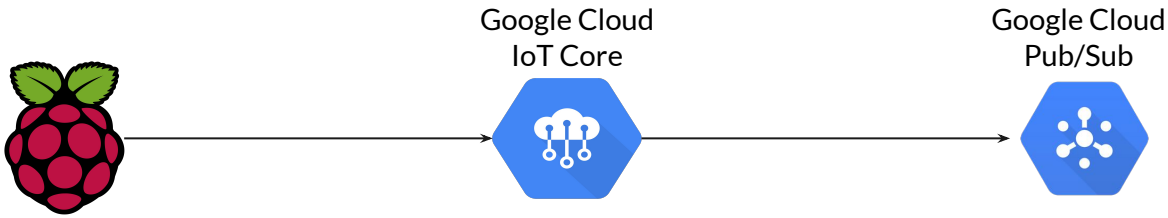


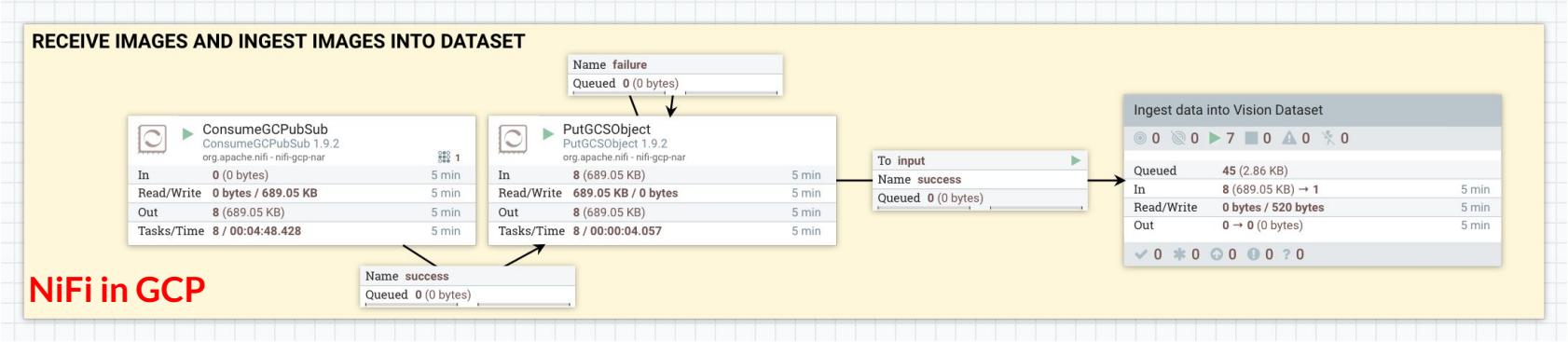
Take pictures and send over MQTT




```
pi@raspberrypi:~/aceu19/raspberry $ cat takePicture.py
from picamera import PiCamera
from gpiozero import LED

imagePath = '/home/pi/aceu19/raspberry/pictures/picture.jpg'

amber = LED(27)

amber.on()

camera = PiCamera()
camera.rotation = 180
camera.capture(imagePath, quality = 10)

amber.off()
```

Ingest images into the dataset





Google Cloud Platform aceu19

Vision

Dashboard

Datasets

Models

aceu19 BETA

LABEL STATS

EXPORT DATA

IMPORT | IMAGES | TRAIN | EVALUATE | TEST & USE

Single-Label Classification

All images	564
Labelled	517
Unlabelled	47

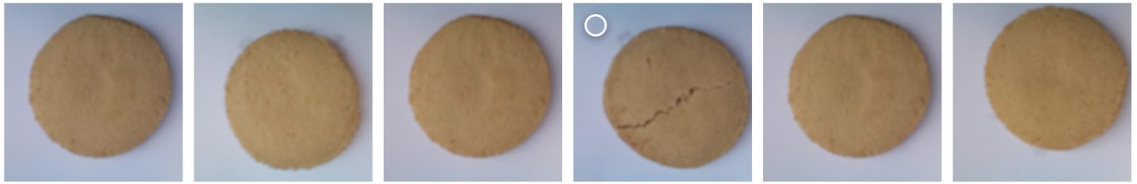
Filter labels

NOK	135
OK	382

ADD NEW LABEL

Filter images

OK(1) OK(1) OK(1) OK(1) OK(1)

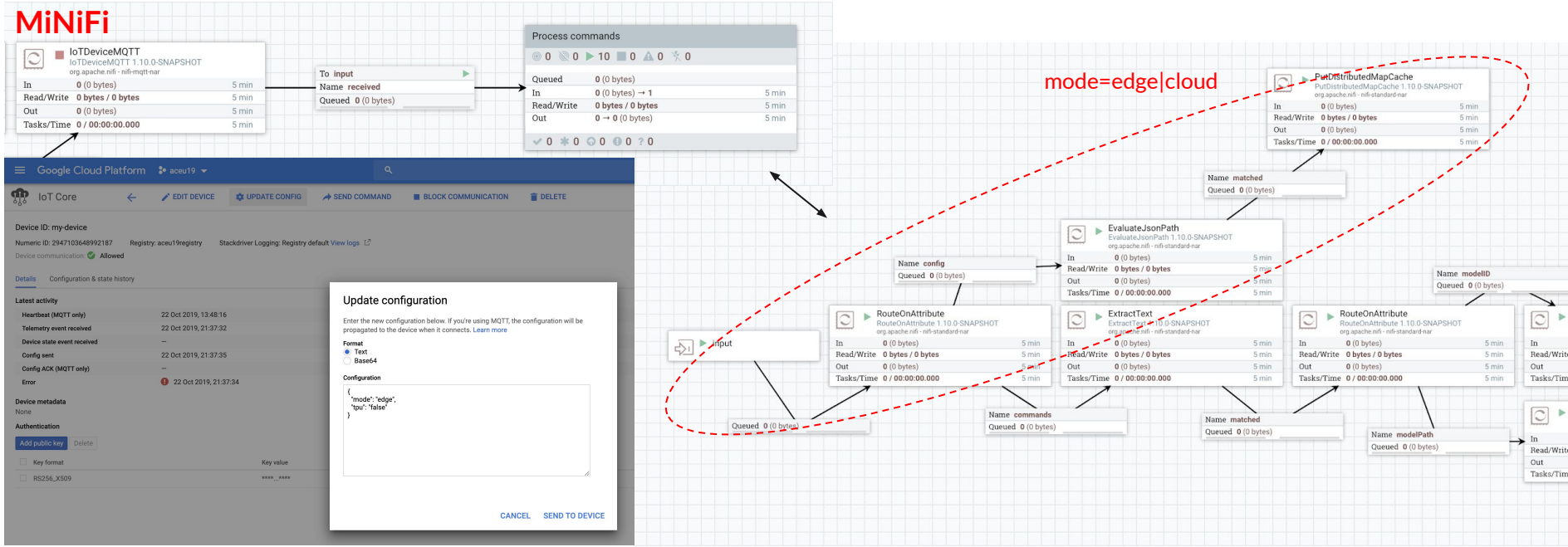


OK(1) OK(1) OK(1) NOK(1) OK(1)

NOK(1) OK(1) OK(1) OK(1) NOK(1) NOK(1)

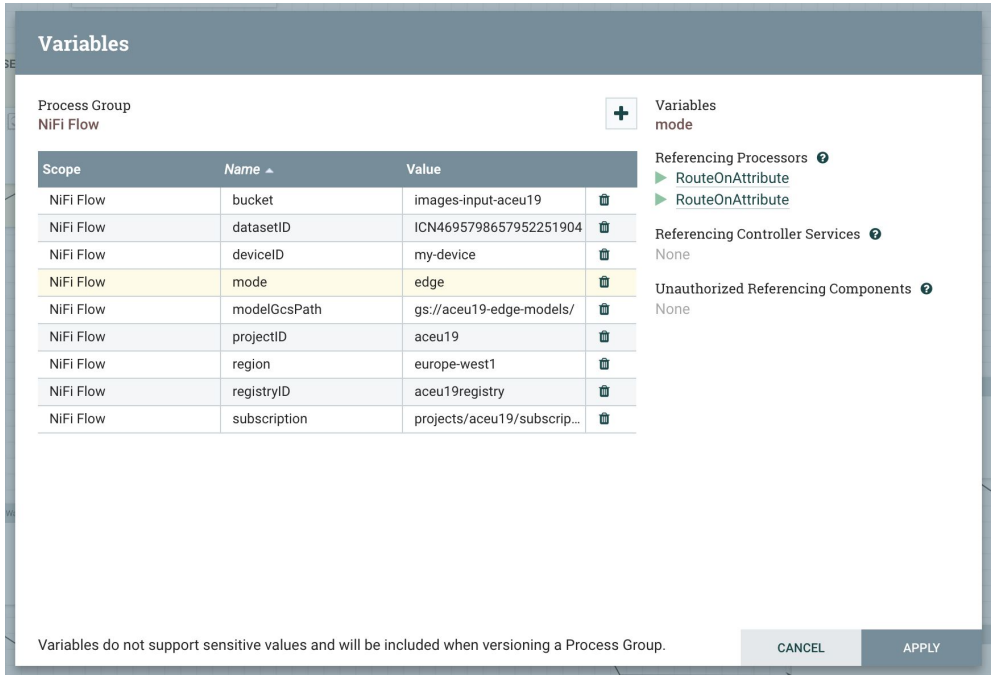
Configure the device

Set device mode: edge/cloud



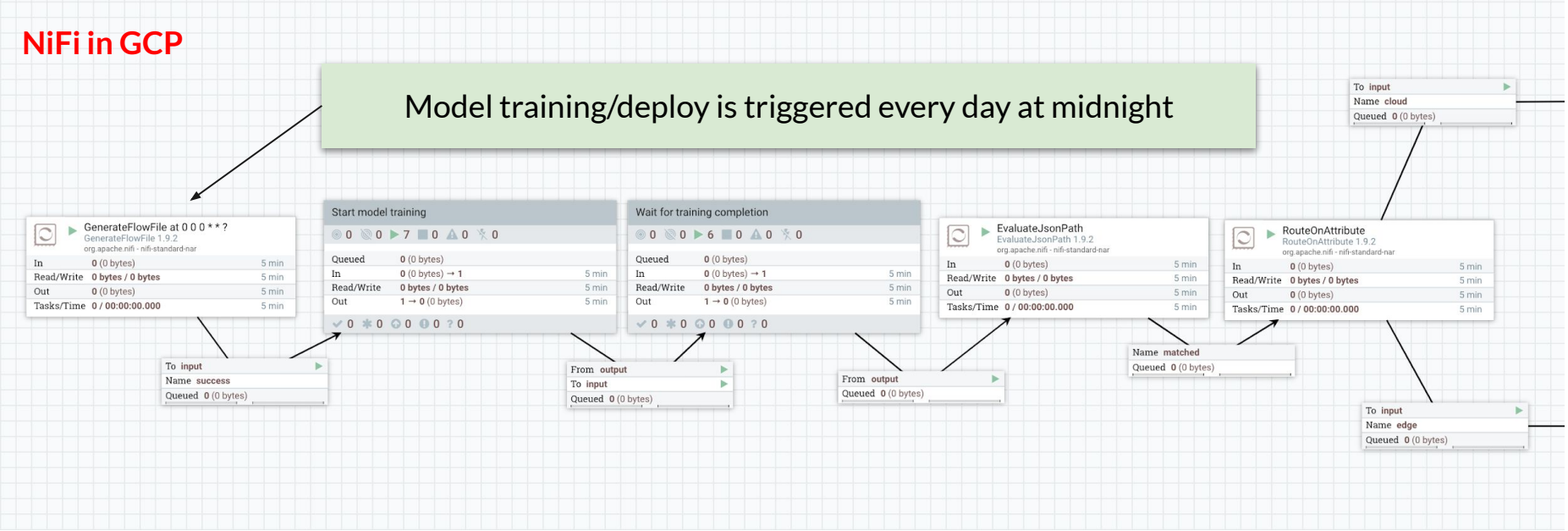
Configure NiFi running on GCP

Variables

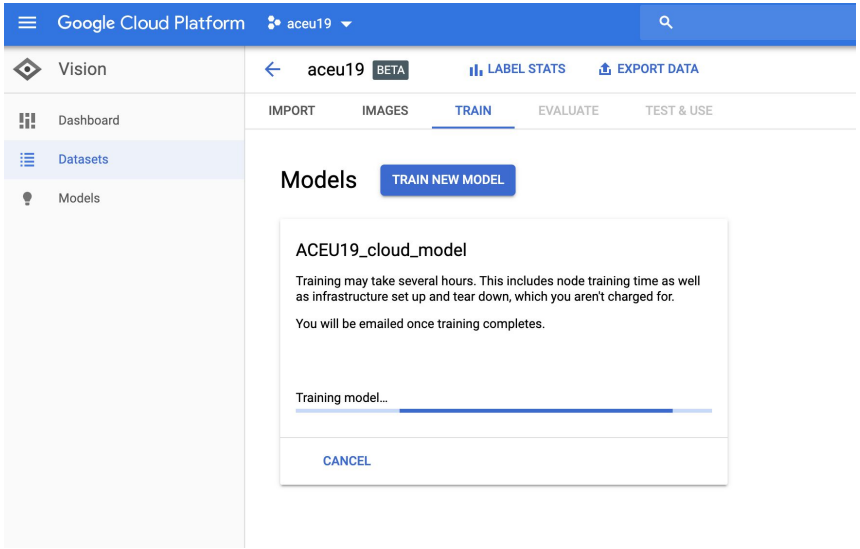

Variables
Process Group
NiFi Flow
Scope
Name
Value
NiFi Flow bucket images-input-aceu19
NiFi Flow datasetID ICN4695798657952251904
NiFi Flow deviceID my-device
NiFi Flow mode edge
NiFi Flow modelGcsPath gs://aceu19-edge-models/
NiFi Flow projectID aceu19
NiFi Flow region europe-west1
NiFi Flow registryID aceu19registry
NiFi Flow subscription projects/aceu19/subscrip...
Variables
mode
Referencing Processors
RouteOnAttribute
RouteOnAttribute
Referencing Controller Services
None
Unauthorized Referencing Components
None
Variables do not support sensitive values and will be included when versioning a Process Group.
CANCEL
APPLY

Model training

Automate model training



Automate model training



View as: original


```
{
  "name": "projects/294759539727/locations/us-central1/operations/ICN7918570593056718848",
  "metadata": {
    "@type": "type.googleapis.com/google.cloud.automl.v1beta1.OperationMetadata",
    "createTime": "2019-10-19T13:26:16.408203Z",
    "updateTime": "2019-10-19T13:44:51.977161Z",
    "createModelDetails": {}
  },
  "done": true,
  "response": {
    "@type": "type.googleapis.com/google.cloud.automl.v1beta1.Model",
    "name": "projects/294759539727/locations/us-central1/models/ICN4037727754588782592"
  }
}
```

Google Cloud Platform aceu19

Vision

Dashboard

Datasets

Models

aceu19 BETA

LABEL STATS

EXPORT DATA

IMPORT IMAGES TRAIN EVALUATE TEST & USE

Single-Label Classification

Model
ACEU19_cloud_model

Confidence threshold 0.85

Filter labels

All labels

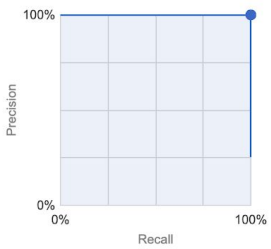
NOK

OK

NOK

Total images	497
Test items	0
Precision	100%
Recall	100%

Use the slider to see which confidence threshold works best for your model on the precision-recall tradeoff curve.
Learn more about these metrics and graphs.



All test images are evaluated at the time of training. If you modify your dataset after training, these results will not be accurate.

True positives

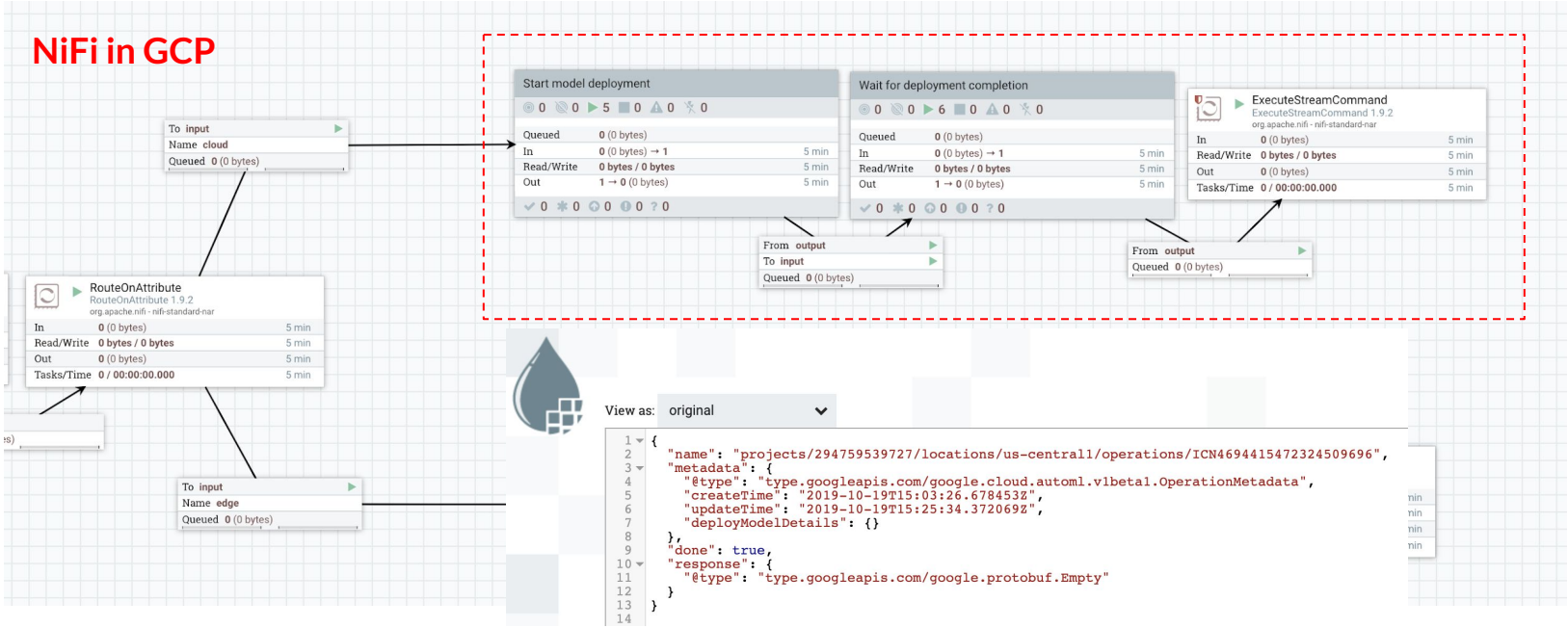
Your model correctly predicted **NOK** on these images

Score: 0.9975636 Score: 0.9988317 Score: 0.9988655 Score: 0.99907696 Score: 0.99919194 Score: 0.99943775

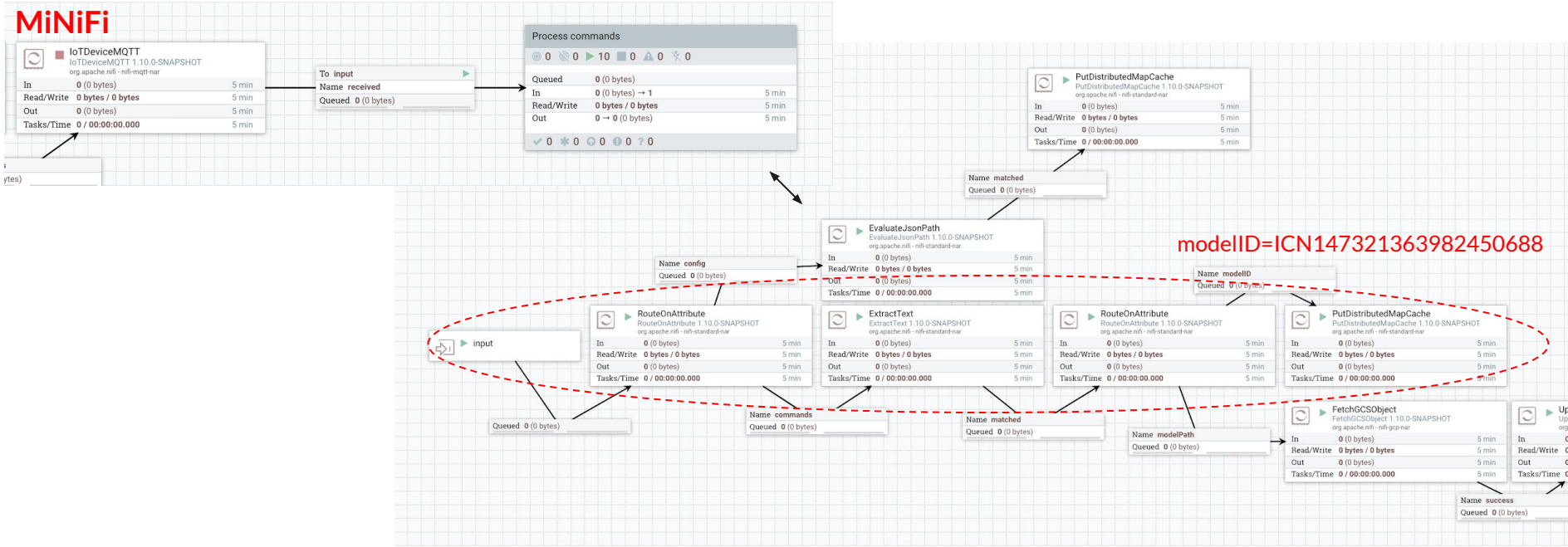
Architecture #1 - Model deployment

TL;DR - the model is running in the cloud

Cloud model deployment



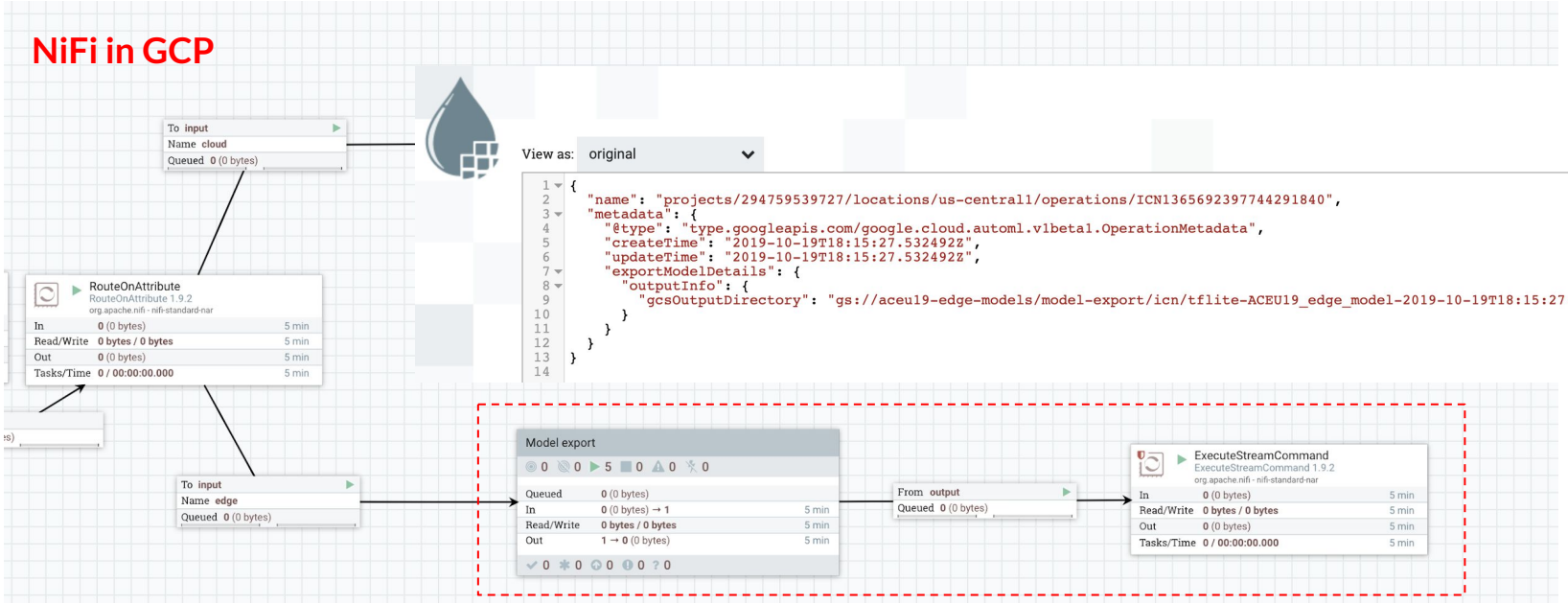
Update on the device



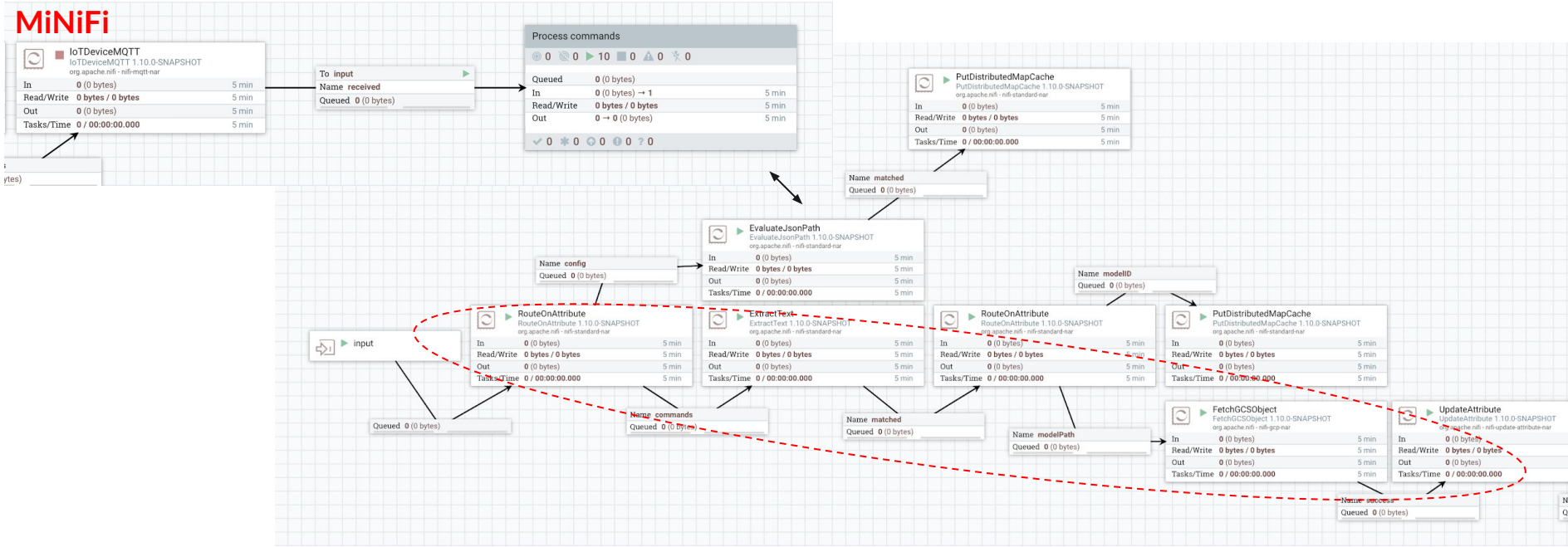
Architecture #2 - Model export

TL;DR - the model is running on the edge

Edge model export into Google Cloud Storage

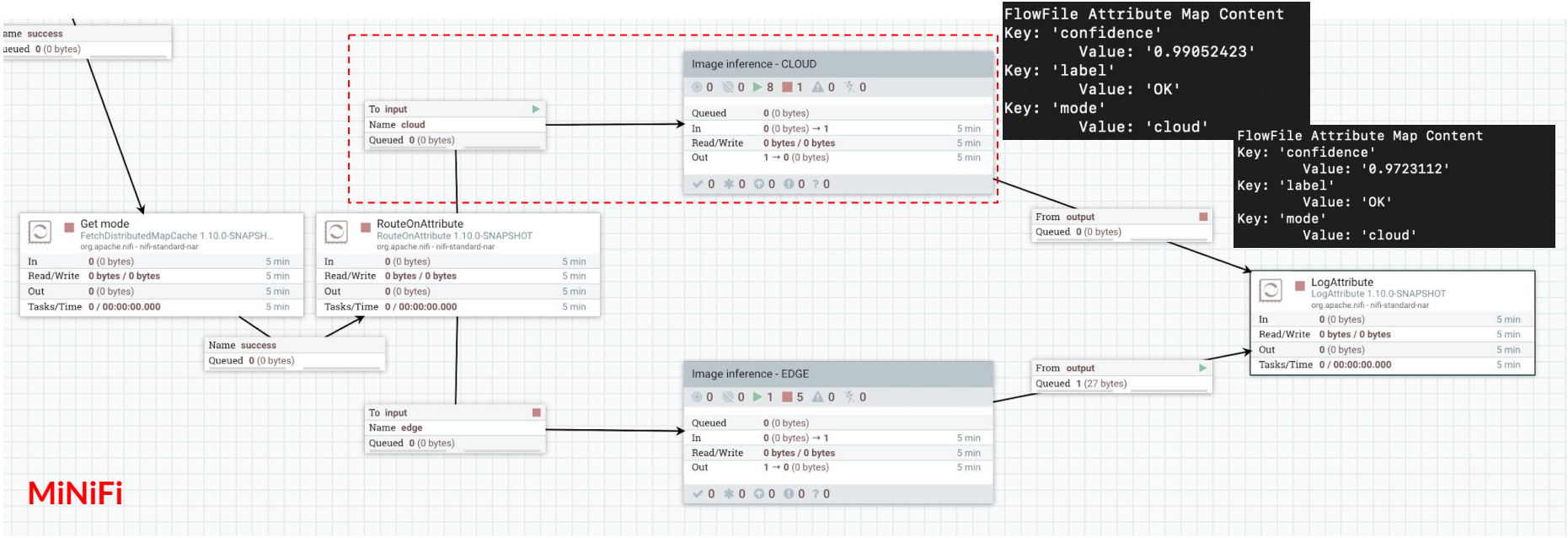


Update on the device

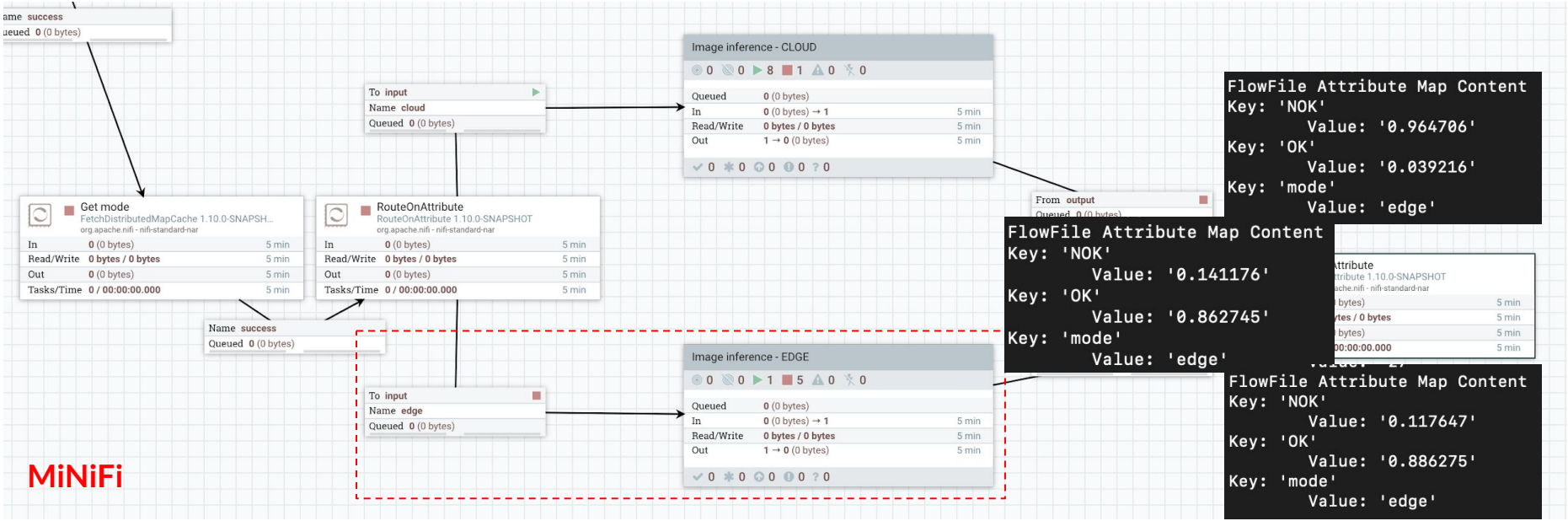


Inference

Inference with model in the cloud



Inference with model on the edge



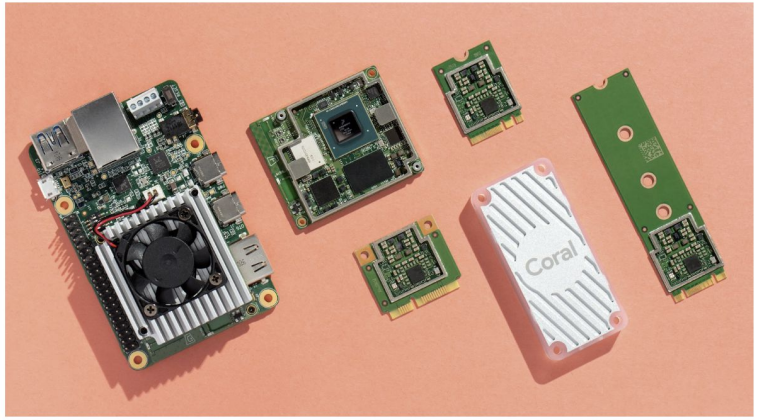
Boosting your edge device with Google's TPU

Coral moves out of beta

Tuesday, October 22, 2019

https://developers.googleblog.com/2019/10/coral-moves-out-of-beta.html

Posted by Vikram Tank (Product Manager), Coral Team



Google Cloud

https://coral.ai/

Processing & Inference time

Architecture	Processing time	Inference time
Model in the cloud	about 6 seconds	about 2.5 seconds
Model on the edge	about 750 milliseconds	about 127 milliseconds
Model on the edge + Coral USB accelerator	about 500 milliseconds	about 9 milliseconds

Model optimized for Edge TPU

```
pi@raspberrypi:~/aceu19/raspberry/inference $ time ./inference/bin/python3 classify_image.py --model model.tflite --label labels.txt --input image.jpg
INFO: Initialized TensorFlow Lite runtime.
----INFERENCE TIME----
Note: The first inference on Edge TPU is slow because it includes loading the model into Edge TPU memory.
136.8ms
127.4ms
127.4ms
127.7ms
127.3ms
-------RESULTS--------
NOK: 0.66797
```

Model non optimized for Edge TPU

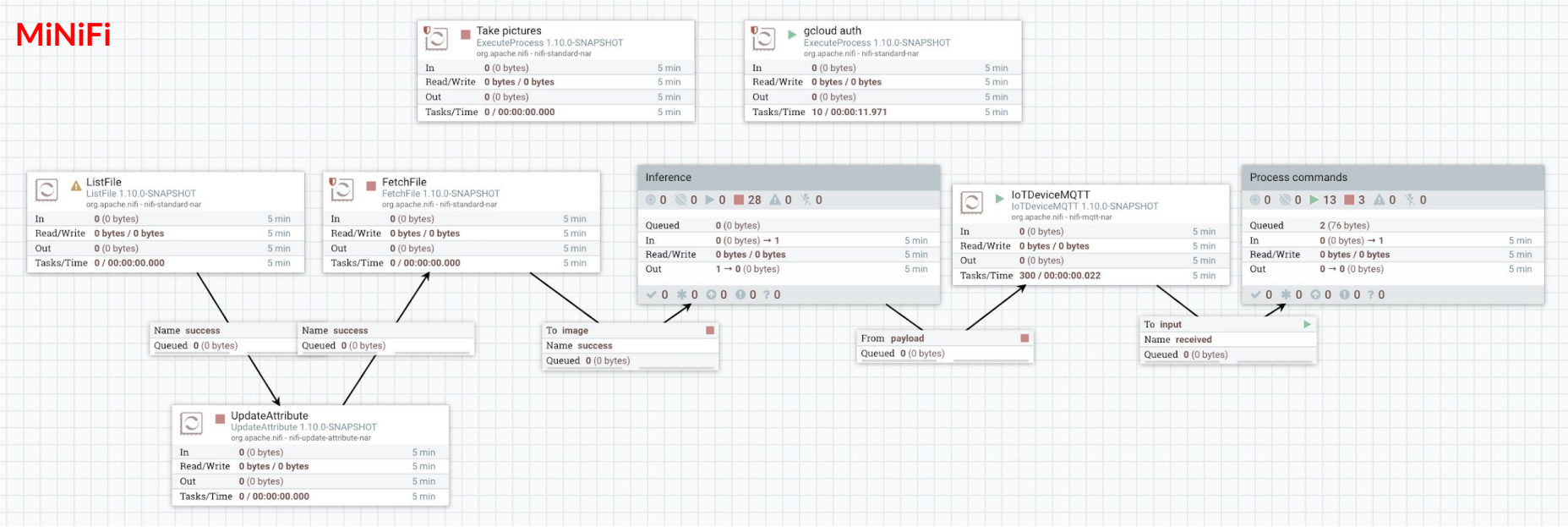
```
pi@raspberrypi:~/aceu19/raspberry/inference $ time ./inference/bin/python3 classify_image.py --model edgetpu_model.tflite --label labels.txt --input image.jpg
INFO: Initialized TensorFlow Lite runtime.
----INFERENCE TIME----
Note: The first inference on Edge TPU is slow because it includes loading the model into Edge TPU memory.
32.9ms
9.2ms
9.2ms
9.0ms
9.1ms
-------RESULTS--------
NOK: 0.67969
```

Model optimized for Edge TPU

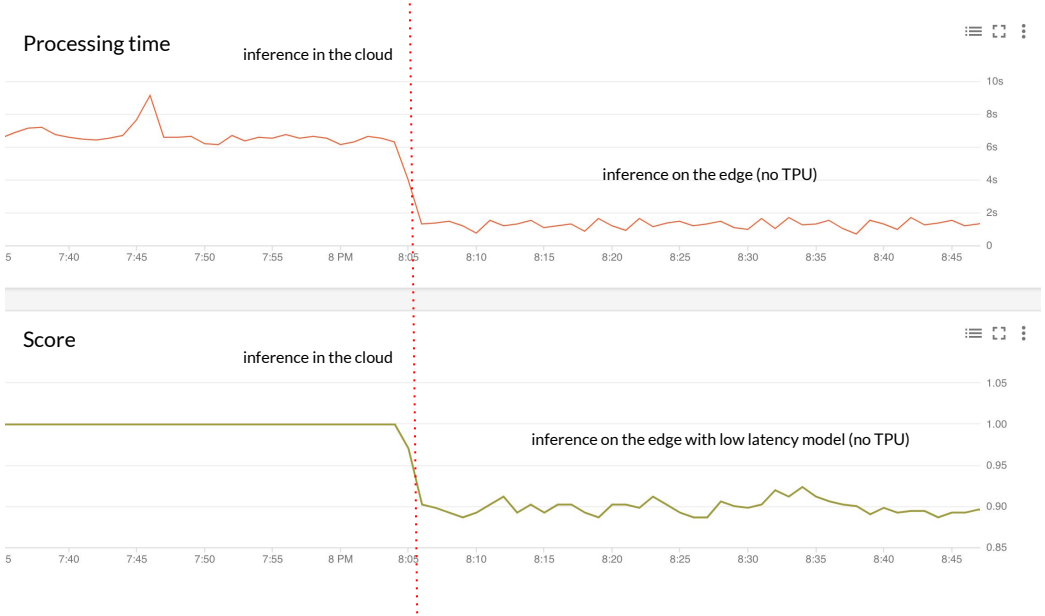
Next steps

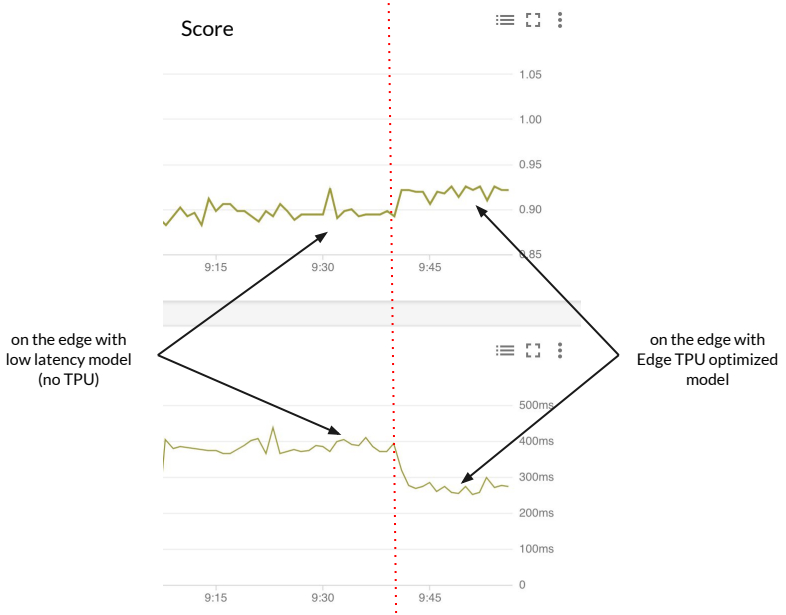
- Auto-labeling when confidence is over a given threshold (example: 0.90)
  - Will drastically reduce human effort to label newly captured data
- Send inference results along with pictures
  - Allow performance monitoring over time, detect outliers and inference performance

Sending inference results along with pictures



Monitoring dashboards in Stackdriver





Analytics in BigQuery

Google Cloud Platform aceu19

BigQuery FEATURES & INFO SHORTCUTS

Query history
Saved queries
Job history
Transfers
Scheduled queries
BI Engine
Resources + ADD DATA
Search for your tables and data sets
aceu19
monitoring
inference

Query editor

```
SELECT mode, tpu, AVG(duration_ms) as duration_ms, AVG(inference_s)*1000 as inference_ms, AVG(score) as score
FROM monitoring.inference
GROUP BY mode, tpu
```

Run | Save query | Save view | Schedule query | More

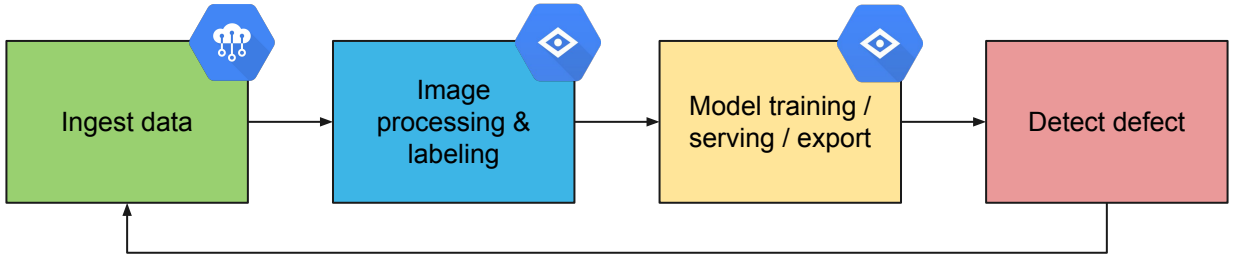
Query results SAVE RESULTS EXPLORE WITH DATA STUDIO

Query complete (1.4 sec elapsed, 24.3 KB processed)

Job information | Results | JSON | Execution details

Row	mode	tpu	duration_ms	inference_ms	score
1	edge	false	1267.0166666666667	385.0352170760269	0.8984375555555554
2	edge	true	1202.5081967213114	273.518443081841	0.9245022950819674

Conclusion - thanks to NiFi and GCP:



- Codeless deployment of customized ML models on the edge
- Feedback loop and continuously updated models
- Processing on the edge and optimization with TF Lite et Coral Edge TPU

THANKS!

Running visual quality inspection at the edge with Apache NiFi & MiNiFi

Pierre Villard - @pvillard31

https://nifi.apache.org
https://github.com/pvillard31/aceu19

APACHE nifi

APACHECON
EUROPE Oct. 22nd - 24th

2019