



OPEN  
AI + DATA  
FORUM



THE LINUX FOUNDATION  
OPEN SOURCE SUMMIT  
NORTH AMERICA

# AI/ML Data Pipeline Processing with Go Microservices Based Solution at the Edge Using Open Source Technology

(AiCSD – AI Connect for Scientific Data -

<https://github.com/intel/AiCSD> )

#ossummit

Neethu Elizabeth Simon  
Samantha Coyle



# Notices and Disclaimers

- Intel technologies may require enabled hardware, software or service activation.
- No product or component can be absolutely secure. Your costs and results may vary.
- Performance results are based on testing as of dates shown in configurations and may not reflect all publicly available updates. See backup for configuration details. Performance varies by use, configuration and other factors. Learn more at [www.Intel.com/PerformanceIndex](http://www.Intel.com/PerformanceIndex)
- Intel is committed to respecting human rights and avoiding complicity in human rights abuses. See Intel's Global Human Rights Principles. Intel's products and software are intended only to be used in applications that do not cause or contribute to a violation of an internationally recognized human right.
- Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy
- No product or component can be absolutely secure. Your costs and results may vary. Results have been estimated or simulated.
- Intel technologies may require enabled hardware, software or service activation
- Intel does not control or audit third-party data. You should consult other sources to evaluate accuracy.
- © Intel Corporation. Intel, the Intel logo, and other Intel marks are trademarks of Intel Corporation or its subsidiaries. Other names and brands may be claimed as the property of others



# Table of Contents

- Introduction
- Project Objective
- Architecture
- UI & Features
- Challenges & Learnings
- Conclusion

# Speakers



**Neethu Elizabeth Simon**

Senior Software Engineer  
Intel Corporation



# Introduction

# Internet of Things (IoT)



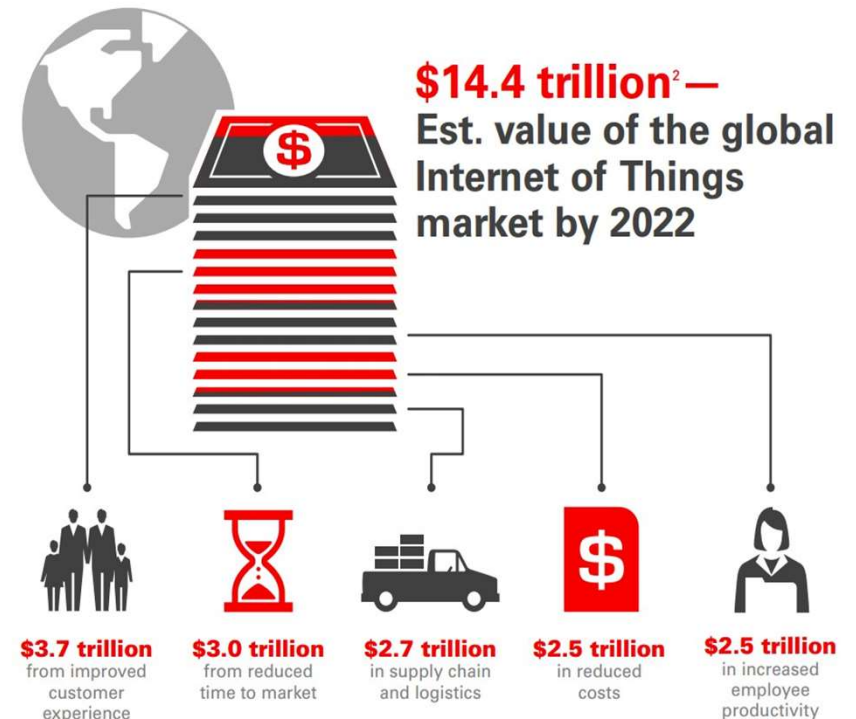
INTERNET OF THINGS IS A GROWING WEB OF  
NUMEROUS DEVICES THAT ARE  
INTERCONNECTED AND INTERACTING



NUMBER OF "THINGS" GETTING INSTALLED  
AND CONNECTED IS GROWING

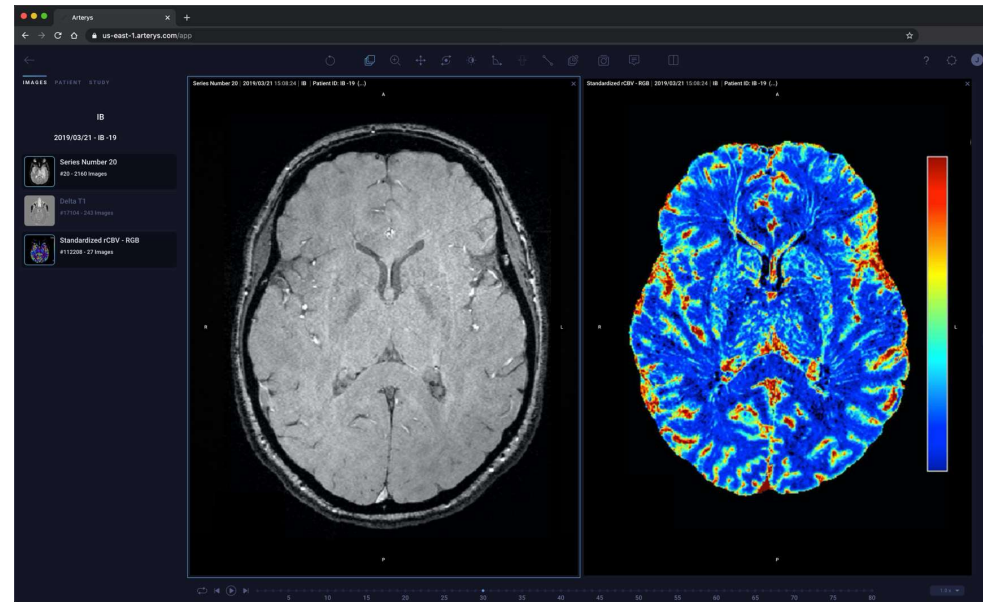
# Growth of IOT

- New and innovative markets
- Increasing industry support
- Strong drivers:
  - Cheaper and faster processors and wireless networks



# Growth of Computer Vision (CV) Applications

- Camera is the ultimate “Thing”
- CV + AI = “Eye of IOT”
- Automated image comparisons adaptable for different use cases & settings

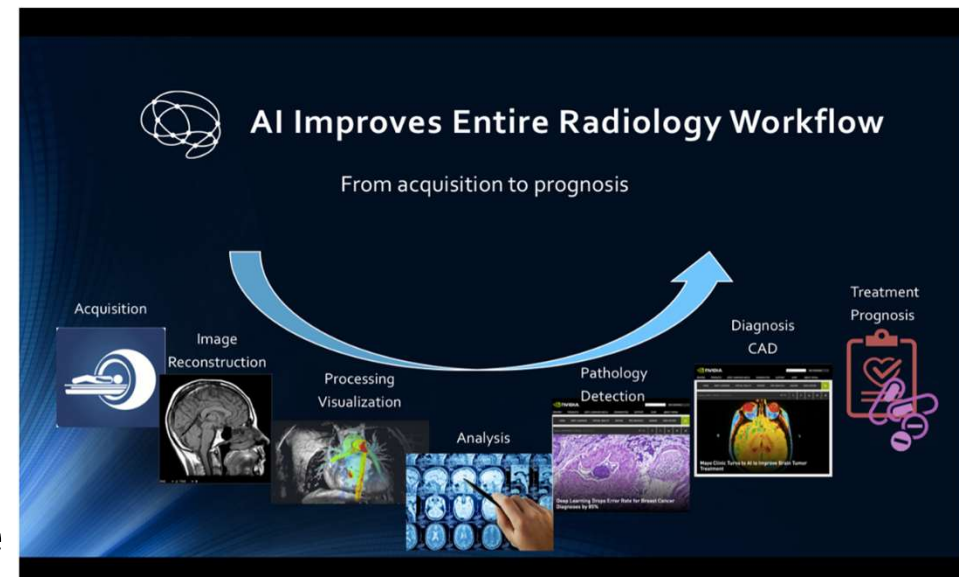


<https://www.arterys.com/>



# Health Care Use Case

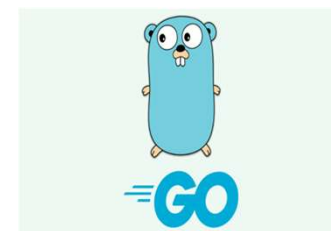
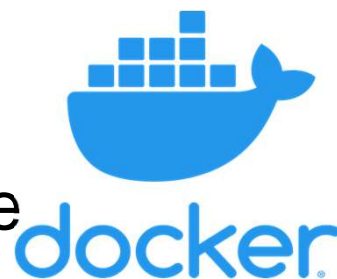
- Focus on health care:
  - Image anomaly detection
  - Edge pipeline enablement
- Deployment pain points
  - Poor AI performance
  - Robustness of architecture
  - Distributed deployment scenario



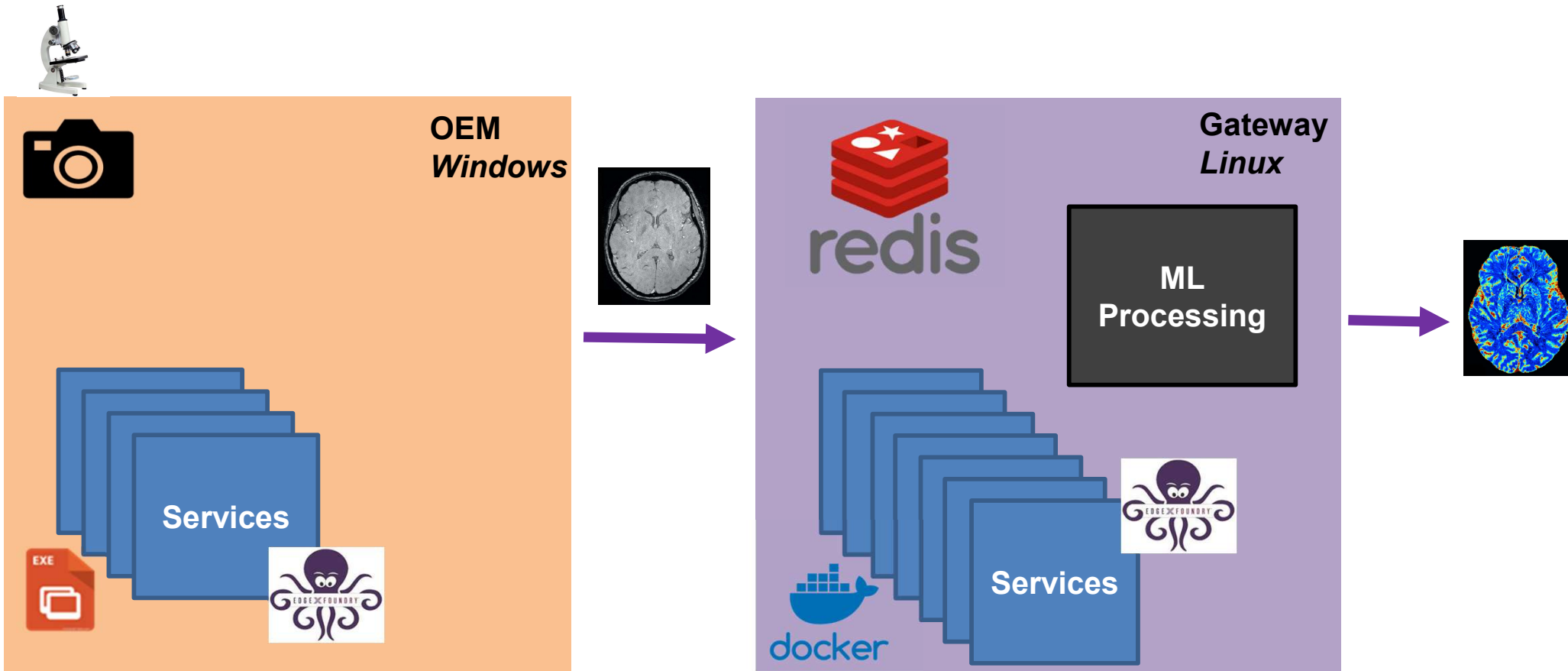
# Project Objective

## Project Objective

- Automated image transfer, processing and comparison
- Go Microservices & Containerized
- Reference Implementation (RI) using Open-Source Software
- AI-Assisted, processing pipeline reducing barrier at the Edge



# Health Care Solution using AI/ML



# Architecture

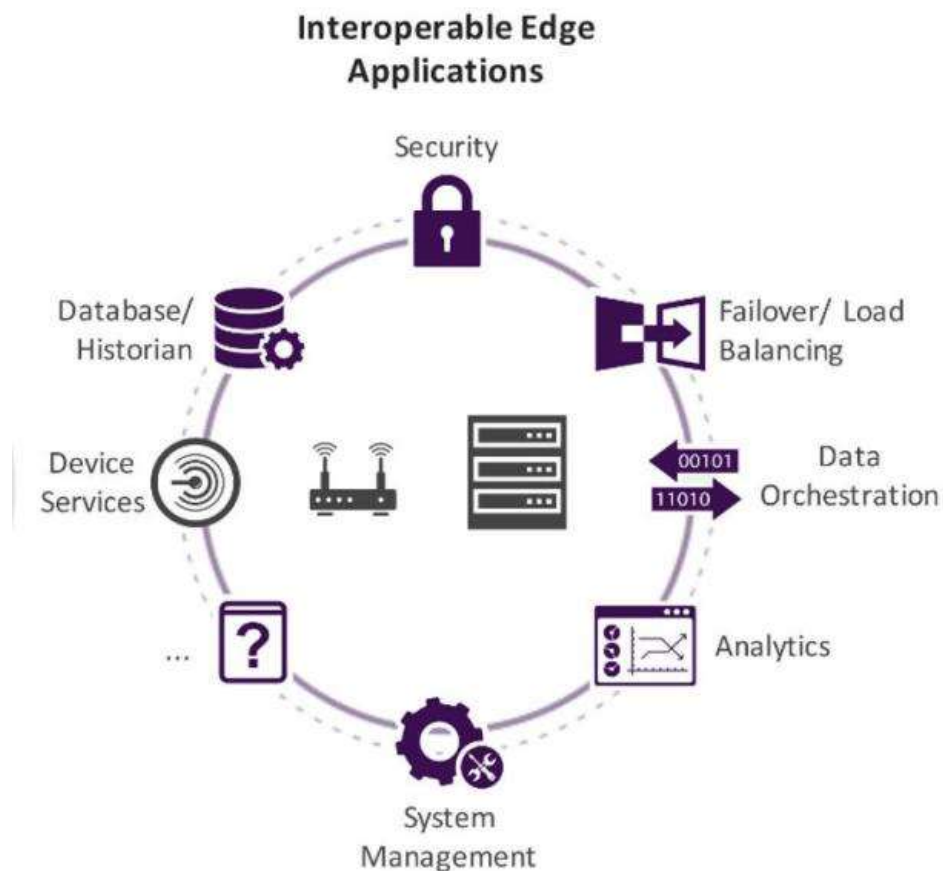
# Terminology

- OEM – Connected to image capturing device
- Gateway – Model inferencing for input images
- Pipeline – AI assisted image pre-processing & inferencing
- Job – Process & Track Images
- Task – Match Jobs to a Pipeline to run

# EdgeX Foundry

## EdgeX:

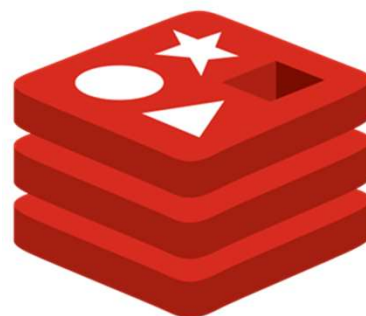
- Hosted by LF Edge
- Framework for IOT edge compute
- Application Service API



# Storage

## Redis:

- In-memory data store
- Seamless EdgeX Foundry integration
- Store Jobs & Tasks



redis

[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)



#osummit



# ML Pipeline Management

## TIBCO Project Air:

- Build, Configure & Modify Pipelines
- Uses EdgeX Foundry



## Intel® Distribution of OpenVINO™ Toolkit

- Open Source Tool for optimized Inferencing on Intel

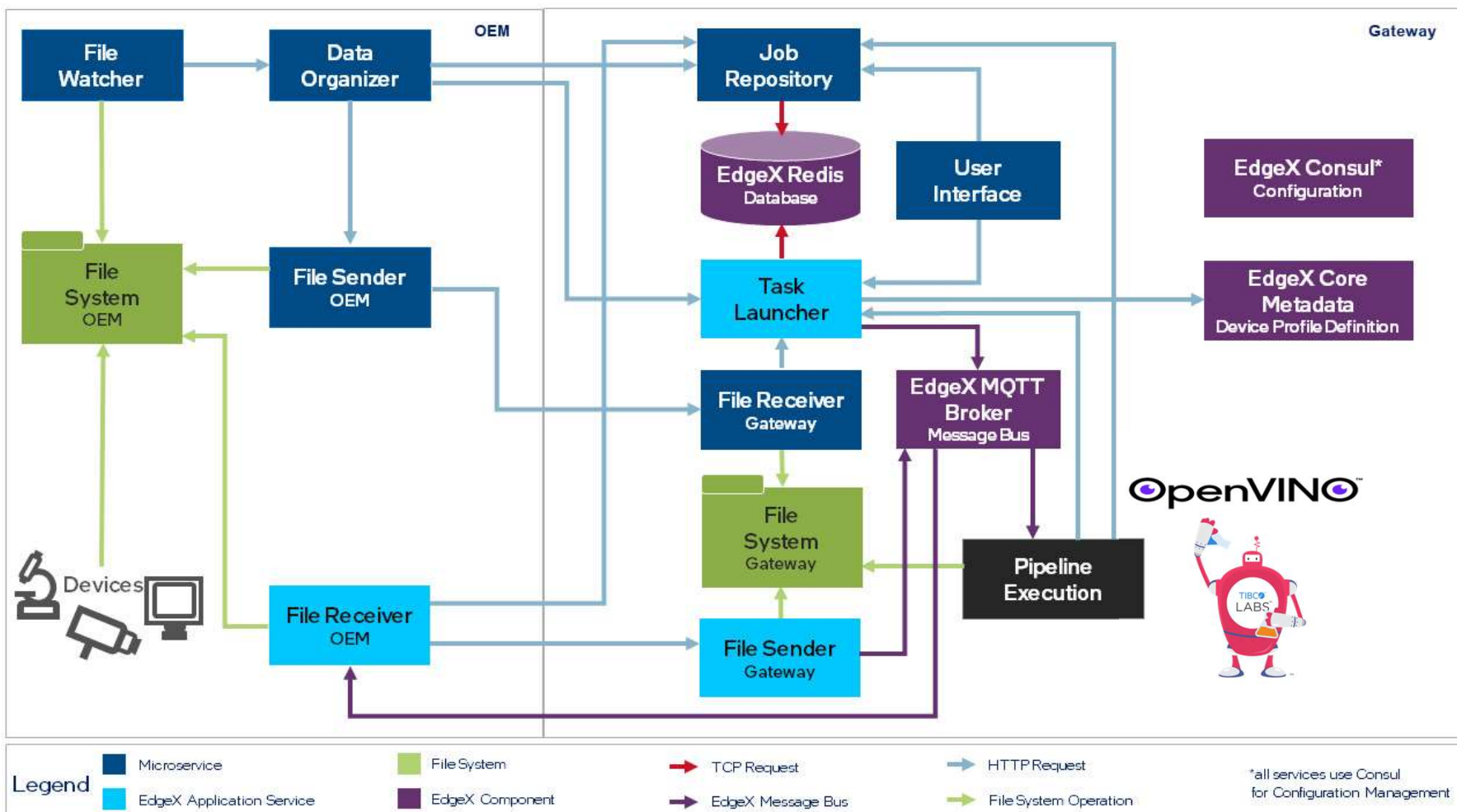


[This Photo](#) by Unknown Author is licensed under [CC BY-SA](#)



#osummit

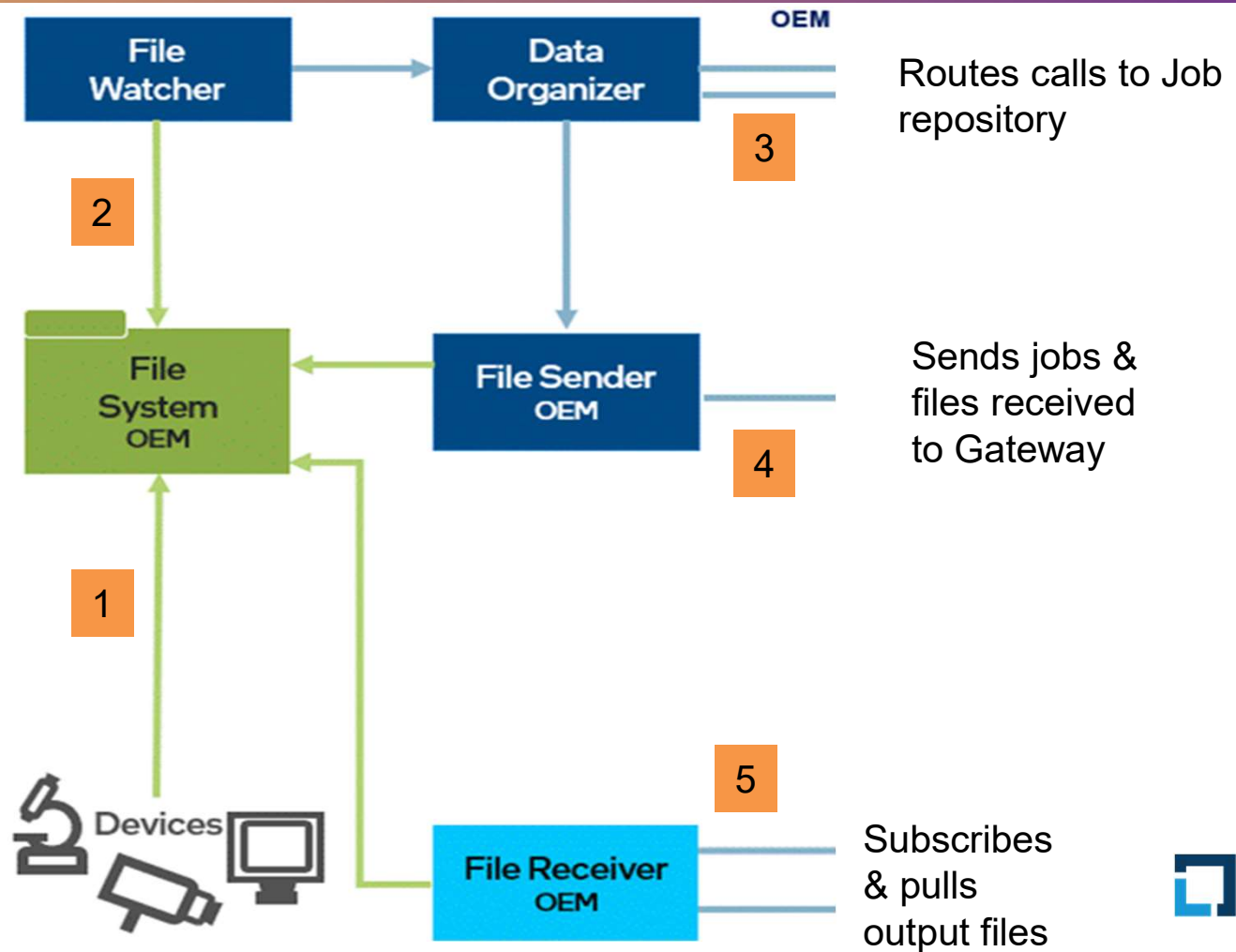
# Architecture



# OEM Microservices

Watch specified folders for images

Send REST call to Data Organizer with Job



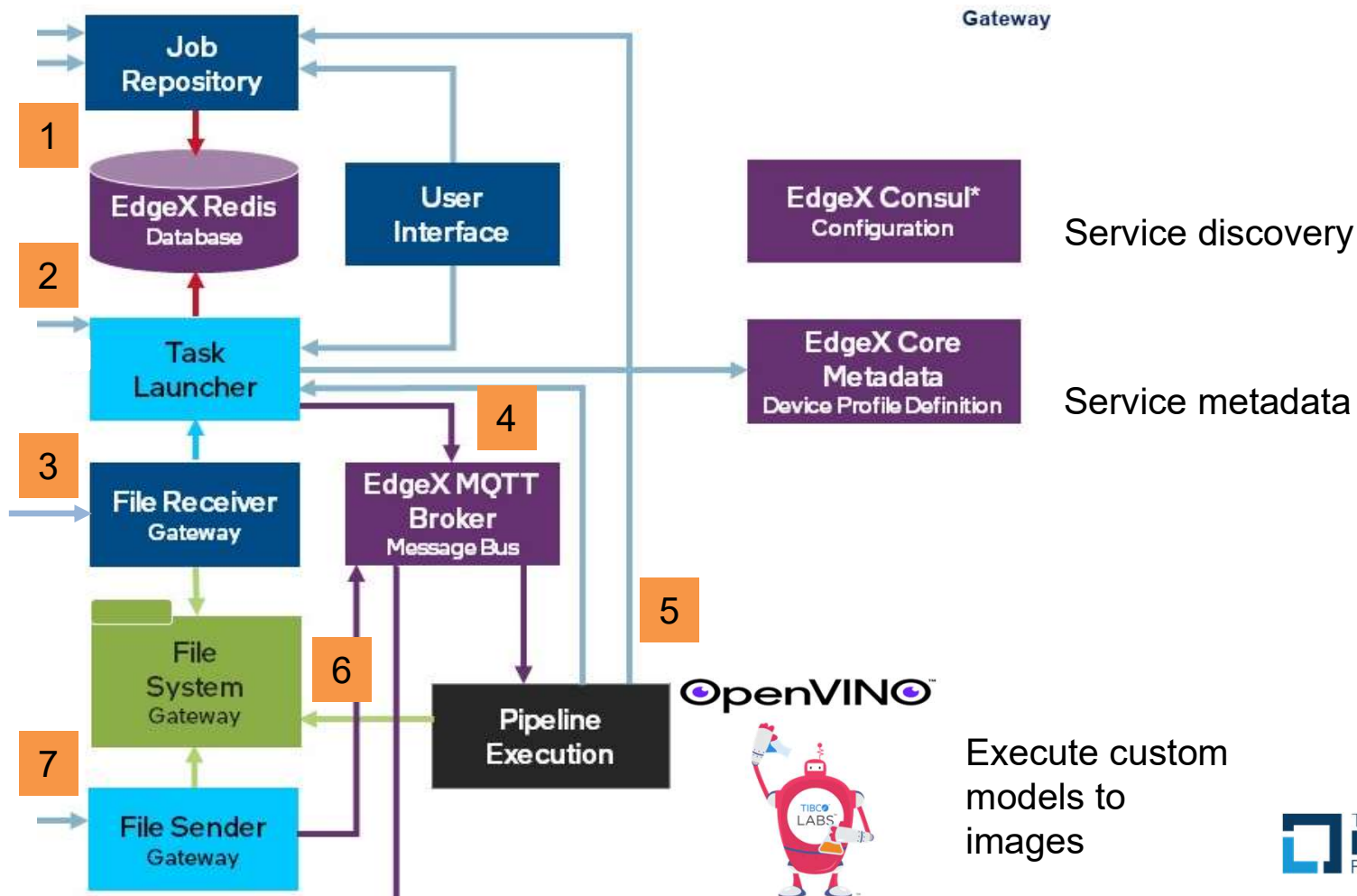
# GW Microservices

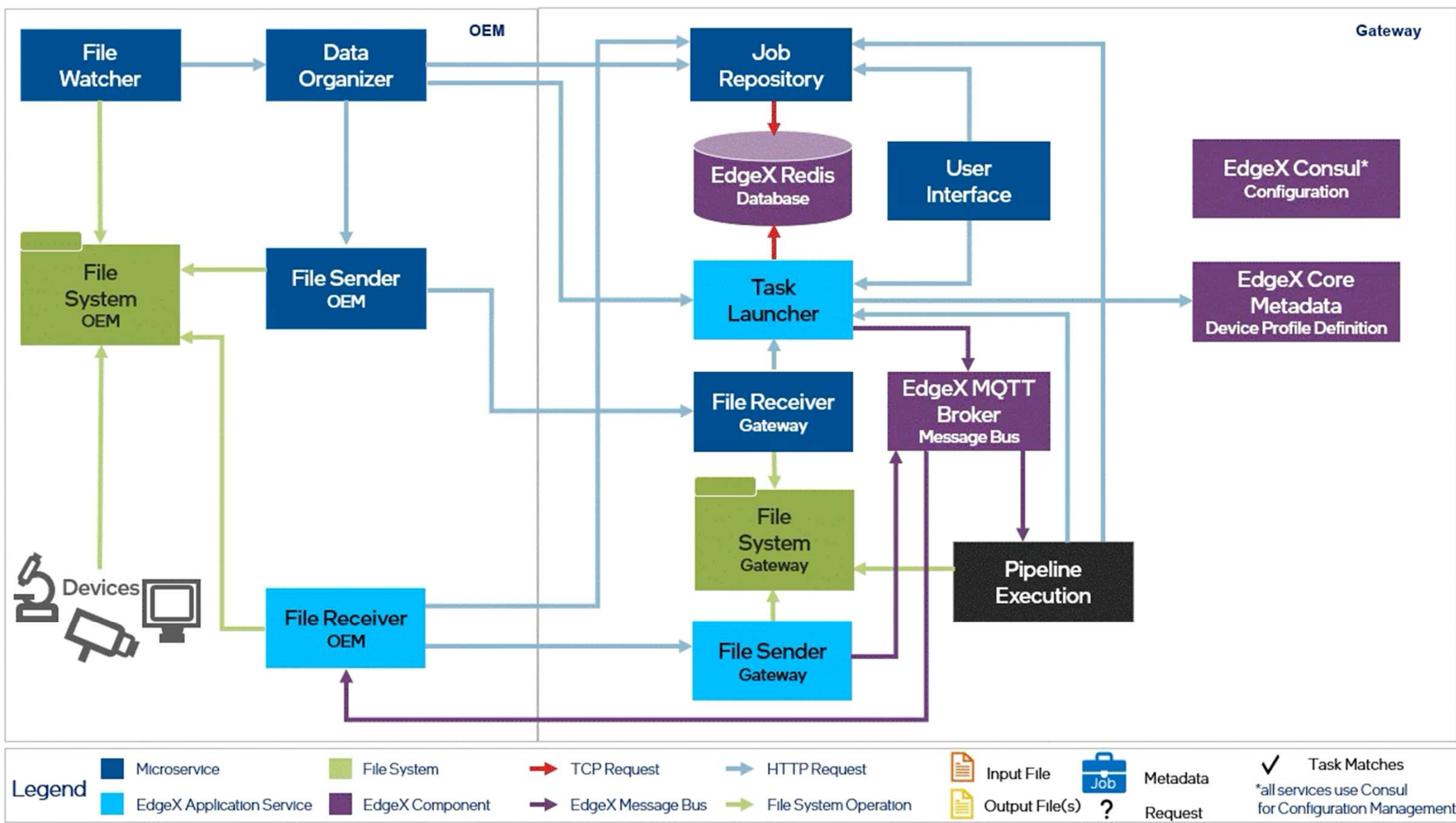
Create Job in Redis

Verifies matching task & send to pipeline

Receiver job & file & write to file system

Publish job & accept request for output files



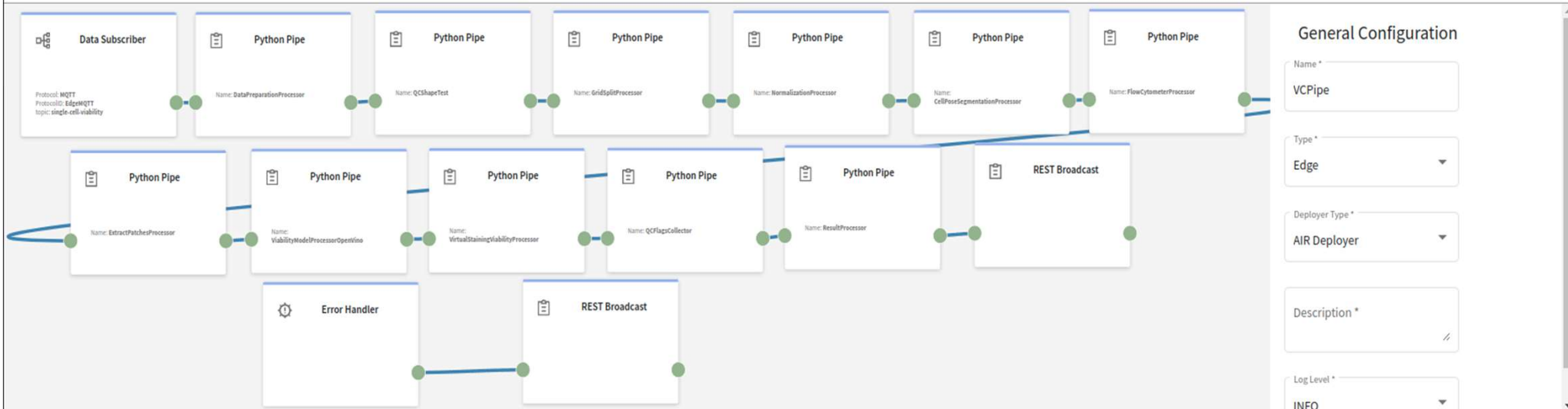


# UI & Features

# Project Air UI

Gateways - Pipelines - ValitaCellGateway

New Pipeline Validate Pipeline Save Pipeline Clear Pipeline Delete Pipeline Print Pipeline **Deploy Pipeline** Undeploy Pipeline



## General Configuration

Name \*  
VCPipe

Type \*  
Edge

Deployer Type \*  
AIR Deployer

Description \*

Log Level \*  
INFO

Id	Name	Pipeline Type	Status	Created	Modified
0x6299b33	VCPipe	Edge	Saved	2022-05-28 18:46:21	2022-05-29 17:35:49

# Project Air Containers

The screenshot shows the Portainer.io web interface. The sidebar on the left contains navigation links: Home, PRIMARY, Dashboard, App Templates, Stacks, Containers, Images, Networks, and Volumes. The main content area is titled "Container list" and shows a table of containers. The table has columns for Name, State, Quick Actions, Stack, and Image. The first container listed is "Air-account\_00001\_0x6299b33" with a state of "running". A tooltip is visible over the container name, displaying "Air-account\_00001\_0x6299b33".

portainer.io

localhost:9000/#!/1/docker/containers

Container list

Containers

admin

my account log out

Containers Columns Settings

Start Stop Kill Restart Pause Resume Remove Add container

Search...

<input type="checkbox"/>	Name	State	Quick Actions	Stack	Image
<input type="checkbox"/>	Air-account_00001_0x6299b33	running		0x6299b33	0x6299b33_0x6299b33
<input type="checkbox"/>	air-app-service-metadata	running		edgex	public.ecr.aws/tibcolabs/labs-air-edgex-app-service-metadata:0.7
<input type="checkbox"/>	air-data-matt-darrah	running		edgex	public.ecr.aws/tibcolabs/labs-air-data-matt-darrah:0.7.0-870



# User Interface

## Angular UI:

- Task Creation
- Job Monitoring
- Features:
  - Observability
  - Internationalization



# Create Task

0.0.0.0:4200/task

intel CREATE/MODIFY TASKS VIEW JOBS OVERVIEW

## Task Management

Configure Tasks

Filter

<input type="checkbox"/>	Description	JobSelector	Pipeline	ResultFileFolder	Model Pipeline	
<input type="checkbox"/>	Generate Output File	matches: "testfile.tiff"	only-file	/tmp/files/output	{ "Brightness": "0", "Resolution": "256" }	<button>Update</button>

Items per page: 5 1 - 1 of 1

**Add Task** Delete Selected

# Select Pipeline

0.0.0.0:4200/task/add

intel CREATE/MODIFY TASKS VIEW JOBS DASHBOARDS

### Add Task

Description \*  
Generate Result

Pipeline \*  
OnlyFile: Pipeline that generates only an output file

Job Selector: ☐ matches ☒ contains

Filename \*  
test-image.tiff

Example: test-image1.tiff

Model Parameters  
{"Brightness": "0", "Resolution": "256"}

Enter the model parameters in following format {"parameter": "value" }

Save Cancel

OnlyFile: Pipeline that generates only an output file

MultiFile: Pipeline that generates multiple output files

OnlyResults: Pipeline that generates only results

FileAndResults: Pipeline that generates output file and results

cancer GetiPipeline: Pipeline that calls Geti pipeline for cancer Det...

# Display Task

The screenshot shows the Intel Task Management web interface. At the top, there's a navigation bar with 'intel' logo and tabs for 'CREATE/MODIFY TASKS', 'VIEW JOBS', and 'OVERVIEW'. Below this is a 'Task Management' section with a 'Configure Tasks' sub-header. A filter input field is labeled with a circled '1'. Below the filter is a table of tasks. The table has columns: 'Description', 'JobSelector', 'Pipeline', 'ResultFileFolder', and 'Model Pipeline'. The first row is a header. The second row is 'Generate Result' with 'contains: "Image.tiff"', 'only-results', '/tmp/files/output', and '{"Brightness": "0", "Resolution": "256"}'. The third row is 'Generate Output File' with 'matches: "testfile.tiff"', 'only-file', '/tmp/files/output', and '{"Brightness": "0", "Resolution": "256"}'. Each row has an 'Update' button to its right. The 'Update' button for the third row is labeled with a circled '3'. The table header row is labeled with a circled '2'. At the bottom right, there's a pagination control showing 'Items per page: 5' and '1 - 2 of 2'. At the bottom left, there are 'Add Task' and 'Delete Selected' buttons.

Task Management  
Configure Tasks

Filter

Description	JobSelector	Pipeline	ResultFileFolder	Model Pipeline	
Generate Result	contains: "Image.tiff"	only-results	/tmp/files/output	{"Brightness": "0", "Resolution": "256"}	Update
Generate Output File	matches: "testfile.tiff"	only-file	/tmp/files/output	{"Brightness": "0", "Resolution": "256"}	Update

Items per page: 5 1 - 2 of 2

Add Task Delete Selected







# Display Jobs

## Jobs

View and Manage Jobs

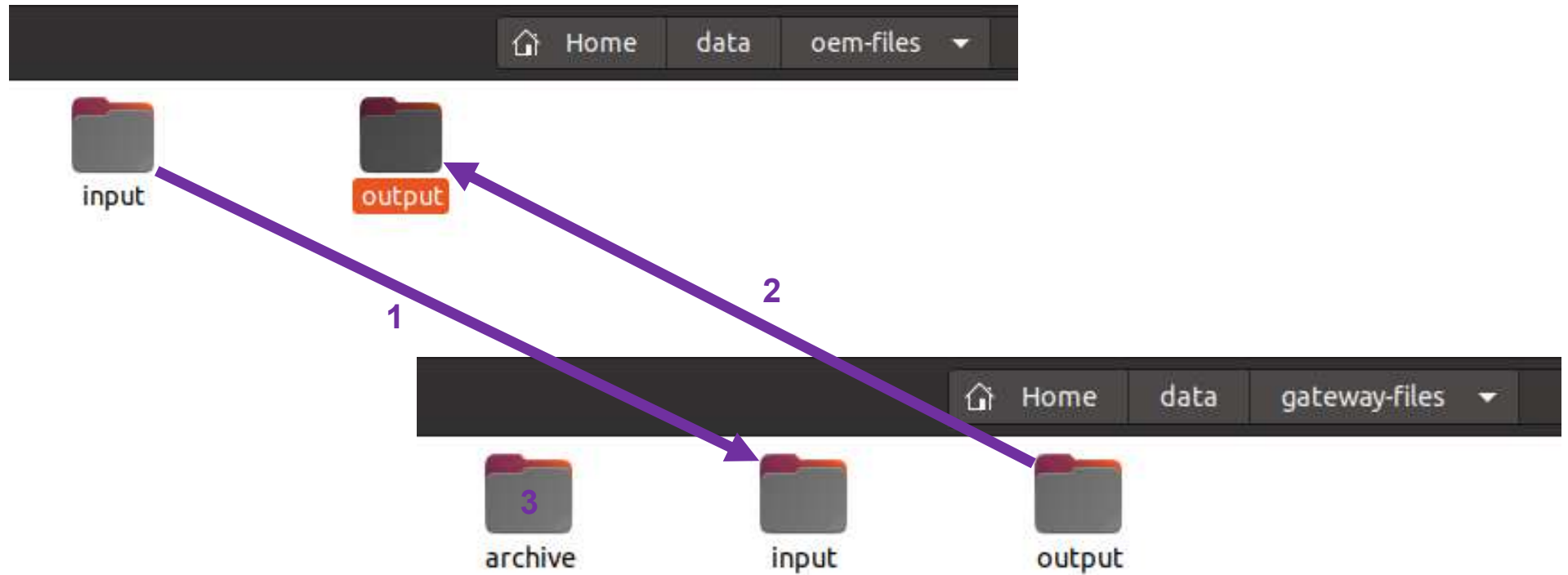
Filter

Expand All Input File Details

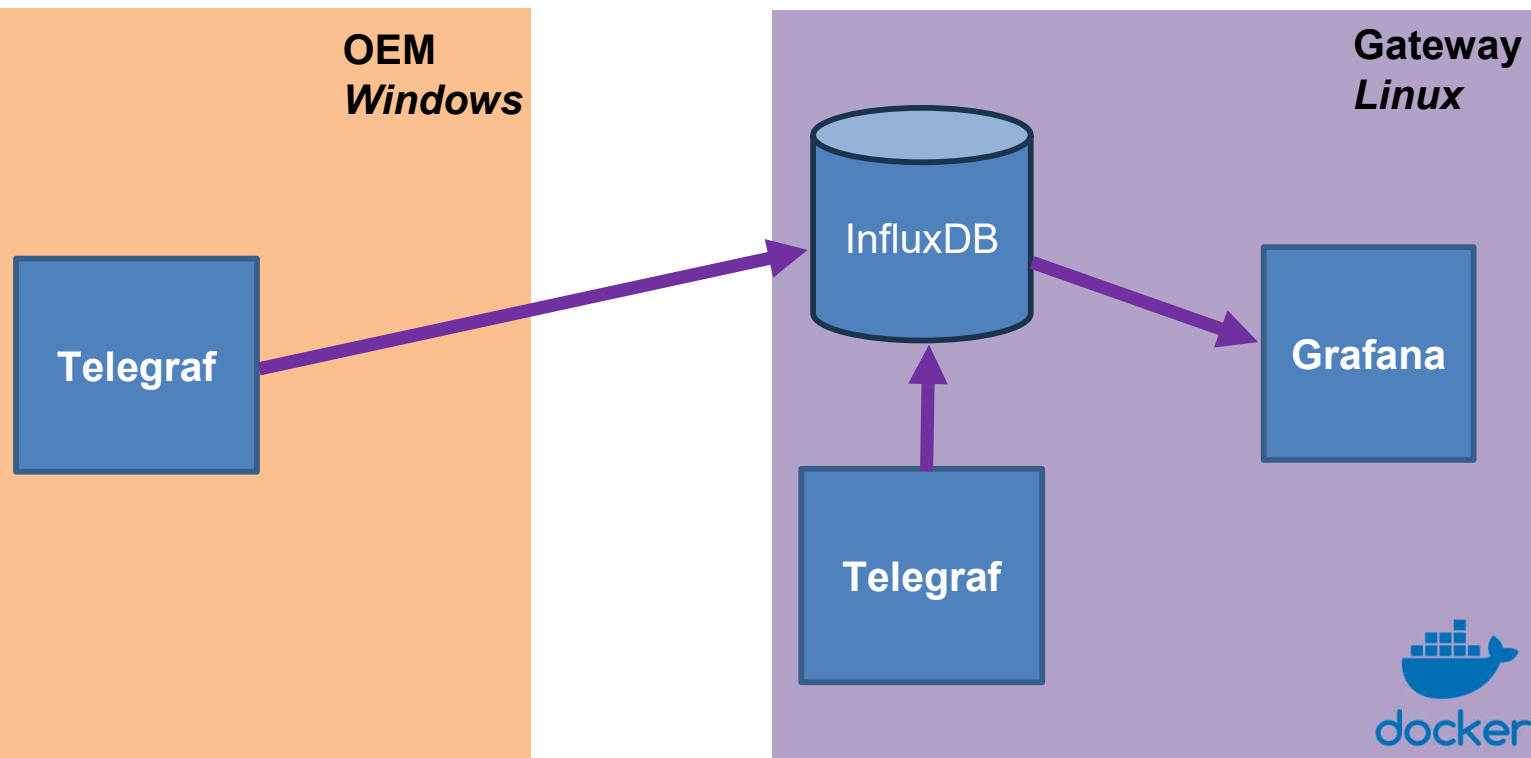
Owner	Job Status	Input Files	Pipeline Status	QCFlags	Output Files	Results	Error Details
None	Complete	 r01c06f06p01-ch1sk1fk1fl1.tiff	PipelineCompletePassed		<a href="#">Oem : [1] Output Files</a>	CellCount, 101	-
None	Complete	 r01c05f01p01-ch1sk1fk1fl1.tiff	PipelineCompleteNone		<a href="#">Oem : [5] Output Files</a>	-	-
None	Pipelineerror	 r01c05f06p01-ch1sk1fk1fl1.tiff	FileNotFound	None	<a href="#">Gateway : [10] Output Files</a>	-	(task-launcher): error accessing file /tmp/files/output/r01c05f01p01-ch1sk1fk1fl1-sim0.tiff: stat /tmp/files/output/r01c05f01p01-ch1sk1fk1fl1-sim0.tiff: no such file or directory
None	Nopipelinefound	 r01c02f02p01-ch1sk1fk1fl1.tiff	-	-	-	-	(data-organizer): no tasks could be matched to the input file name
None	Nopipelinefound	 r01c04f06p01-ch3sk1fk1fl1.tiff	-	-	-	-	(data-organizer): no tasks could be matched to the input file name
None	Complete	 r01c06f05p01-ch1sk1fk1fl1.tiff	PipelineCompletePassed		<a href="#">Oem : [1] Output Files</a>	CellCount, 101	-

Items per page: 10 1 - 6 of 6 |< < > >|

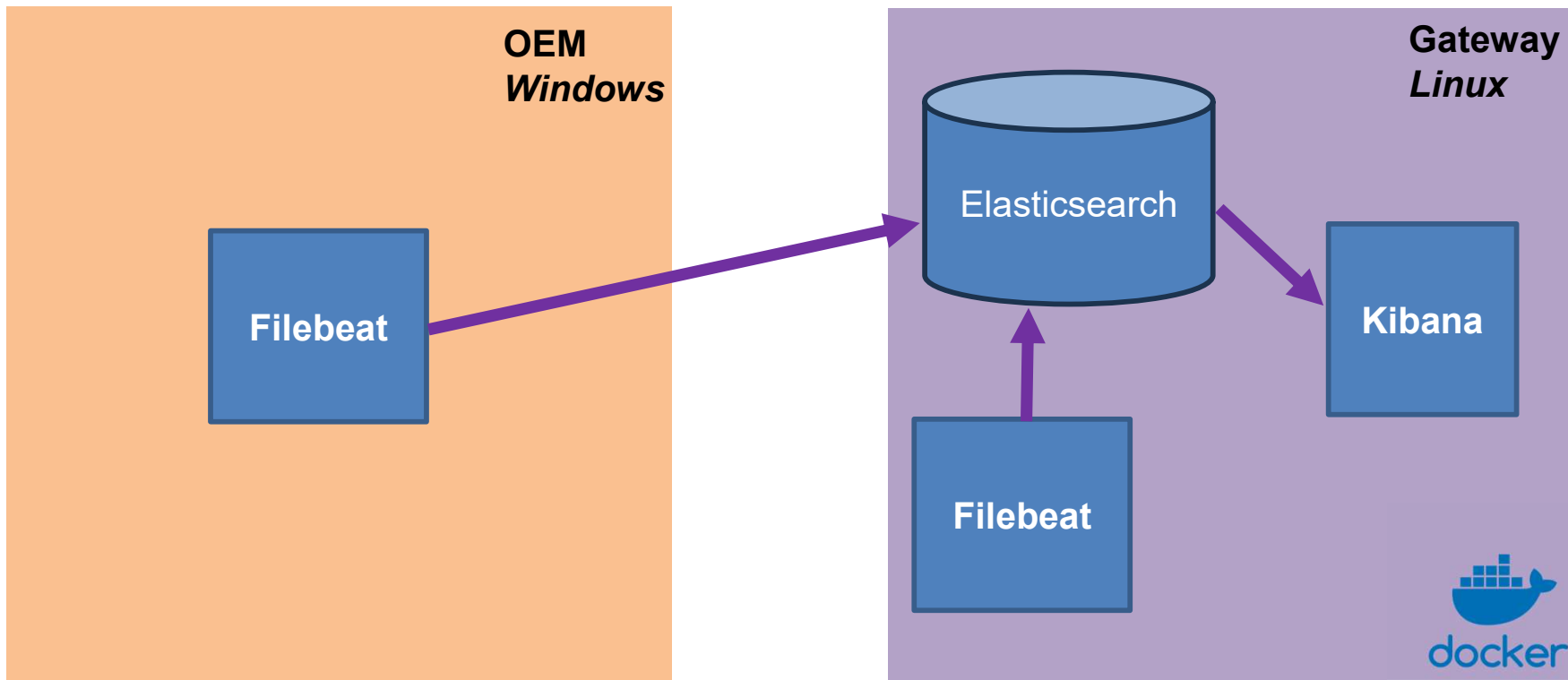
# Input & Output File Folders – OEM & Gateway



# Observability



# Log Analytics





# Internationalization

- Localization of many Job fields to Chinese
- [Go-i18n](#) package
- Update headers:
  - Accept-Language: en-zh



# Integration Testing – Open Source Go Packages

## 1. Integration tests

Test  
Dependencies

Test Cases



Testcontainers



httpexpect

## 2. Test artifacts

Test Report



testreport



#osummit

# Challenges & Learnings

# Distributed Deployment Scenario

- Additional hardware requirements
- Go-MSI package -> Zip folder of .bat files
  - Considerations:
    - Installation process challenges
    - Software requirements
- Different deployment steps per machine

# Distributed Deployment Scenario

- OEM device services:



- Gateway services:



# Distributed Deployment Scenario

## Automation difficulties

- OEM device services:



- Gateway services:

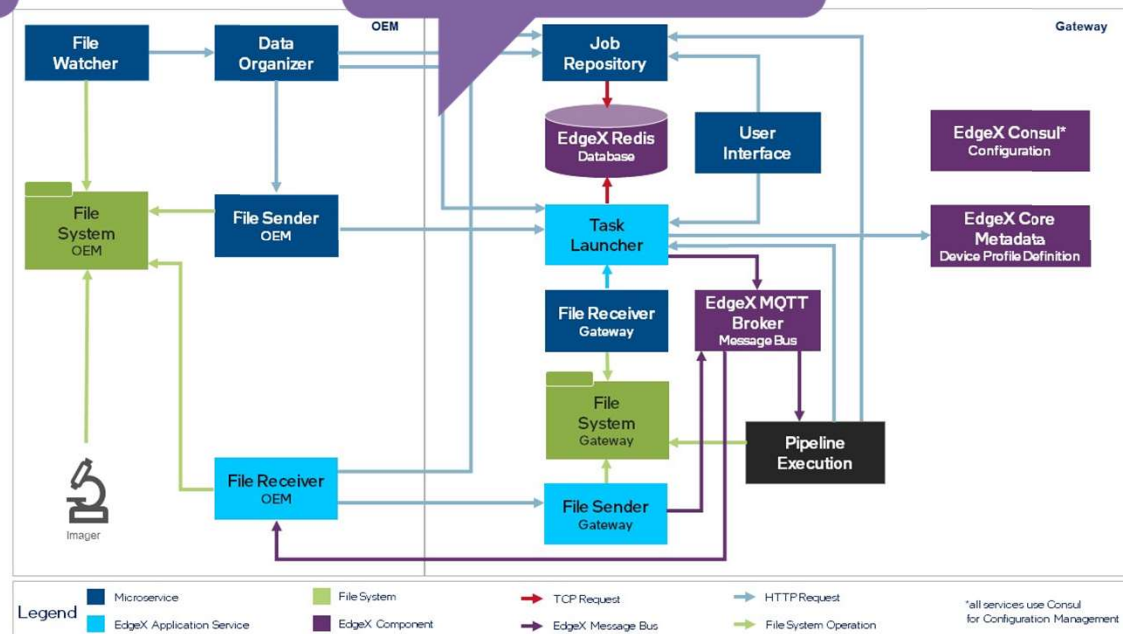


# Service Wait Strategy

Deployment  
time

What ifs?

Integration  
test time



# Wait Strategy Solution

- Implementation layer considerations:
  - Docker VS Go application code layer
- Wait-for-it Go package
  - More challenges 😊
    - Bringing in components into our codebase
    - Minor modification to allow for ease-of-use
    - Open Source PR created, and in-progress
- Consistent service wait strategy





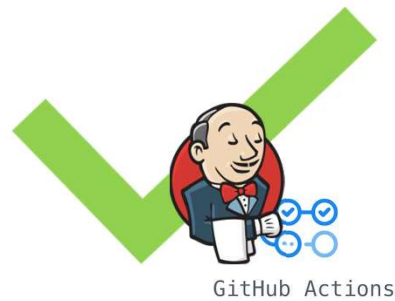
# Idempotency Issues



3 seconds



5 seconds



3 seconds



3 seconds



4 seconds

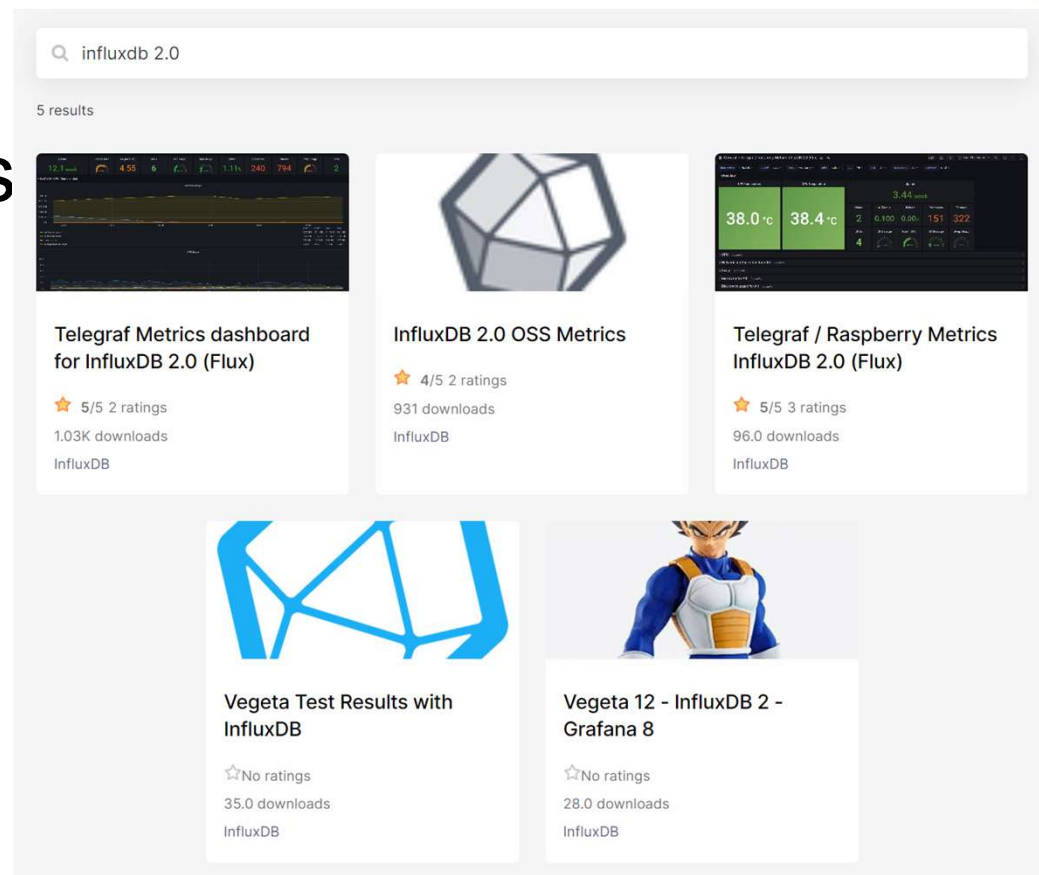
# Observability



## InfluxDB 2.0

## Flux-style queries

## Cutting edge



## Telegraf Metrics dashboard for InfluxDB 2.0 (Flux)

Dashboard for displaying basic host metrics collected by telegraf and stored into the InfluxDB 2.0. Metrics are fetched by flux.

Overview **Revisions** Reviews

### Dashboard revisions

Revision	Description	Created	
1	This is a copy of a dashboard I found here, modified to work with flux syntax. I have no idea if everything is correct, and it could use work.	2022-01-31T01:25:16	<a href="#">Download</a>

## Telegraf Metrics dashboard for InfluxDB 2.0 (Flux)

Dashboard for displaying basic host metrics collected by telegraf and stored into the InfluxDB 2.0. Metrics are fetched by flux.

Overview **Revisions** Reviews

### Dashboard revisions

Revision	Description	Created
1	This is a copy of a dashboard I found here, modified to work with flux syntax. I have no idea if everything is correct, and it could use work.	2022-01-31T01:25:16

[Download](#)

# Open Source Project Shifts

## TIBCO Project Air

TIBCOSoftware / labs-air

Code Issues 5 Pull requests 9 Actions Projects Security Insights

labs-air Public

Watch 12 Fork 17 Star 24

master 10 branches 28 tags

Go to file Add file Code

Gerromie chore(release): 0.8.0-164 ✓ 8bce2bb on Aug 22, 2022 492 commits

.github	fix(cicd): fix issue with npm install and add additional projects	last year
docs-src	Update index.md	last year
docs	fix: v2 endpoints	last year
.commit-template	feat(cicd): introduction of static tools	2 years ago
.gitignore	fix: v2 endpoints	last year
.versionrc.json	chore: change section name to match convention	last year
.whitesource	Add .whitesource configuration file	4 years ago
CHANGELOG.md	chore(release): 0.8.0-164	last year

About

TIBCO LABS™ Project AIR - Documentation

[tibcosoftware.github.io/labs-air/](https://tibcosoftware.github.io/labs-air/)

iot dgraph flogo spotfire edgex

tibco-labs computedb

Readme

BSD-3-Clause license

Activity

24 stars

12 watching

17 forks


Report repository



#ossummit

# Conclusion

## Conclusion

- Replaced Project Air with Open-Source Tool - BentoML
- Released as Open-Source Sample in July
- AiCSD – AI Connect for Scientific Data -  <https://github.com/intel/AiCSD>
- Successfully Deployed customer's lab in Colorado
- [Recorded Webinar on using OpenVINO optimization to unlock AI performance for CellAI toolbox](#)
- Scale to other use cases



# Thank You

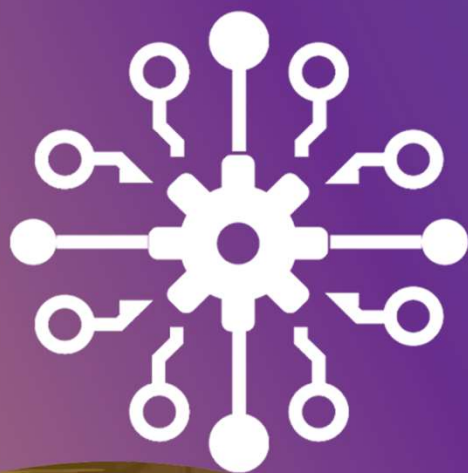
# Questions?



**Neethu Elizabeth Simon**  
Senior Software Engineer  
Intel Corporation







# OPEN AI + DATA FORUM

@



OPEN SOURCE SUMMIT  
NORTH AMERICA

THE LINUX FOUNDATION

