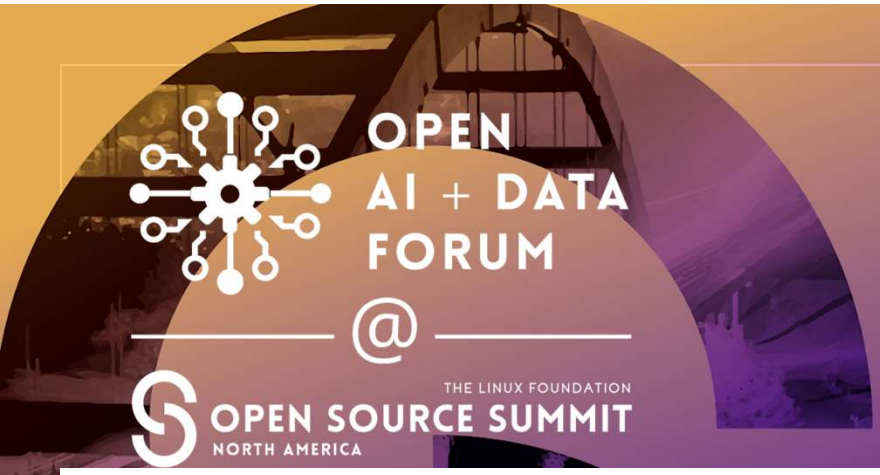


AI/ML Data Pipeline Processing with Go Microservices Based Solution at the Edge Using Open Source Technology

(AiCSD – AI Connect for Scientific Data - https://github.com/intel/AiCSD)

#ossummit

Neethu Elizabeth Simon
Samantha Coyle

THE LINUX FOUNDATION

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
- 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.
-

Table of Contents

Speakers

Neethu Elizabeth Simon
Senior Software Engineer
Intel Corporation

Introduction

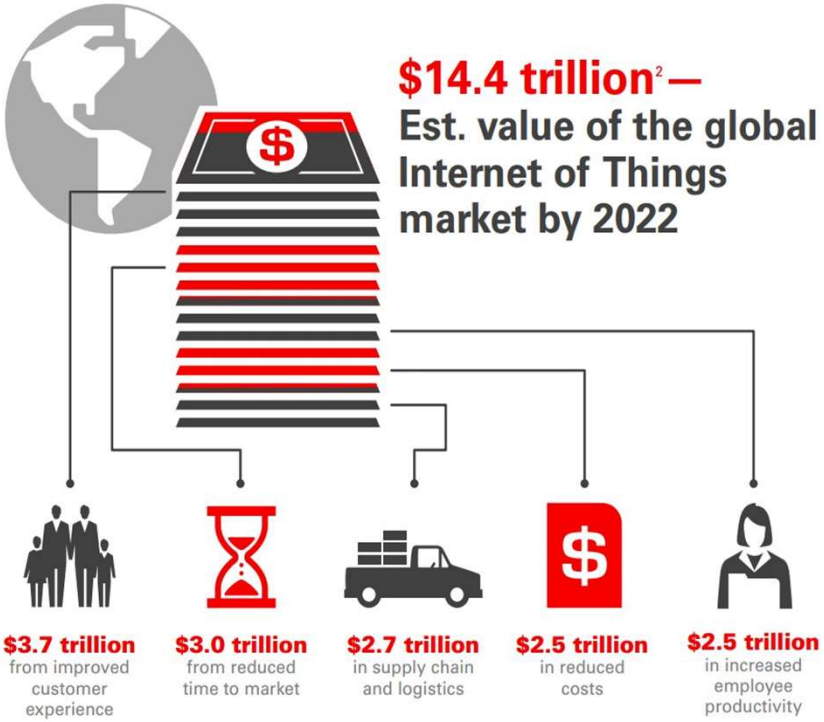
Internet of Things (IOT)

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

NUMBER OF "THINGS" GETTING INSTALLED AND CONNECTED IS GROWING

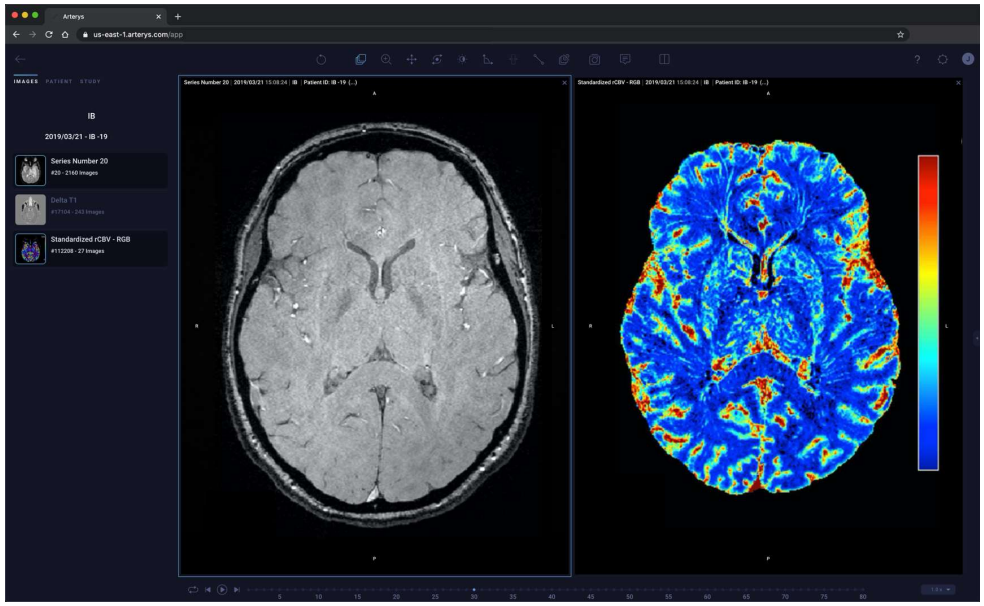
Growth of IOT

- 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

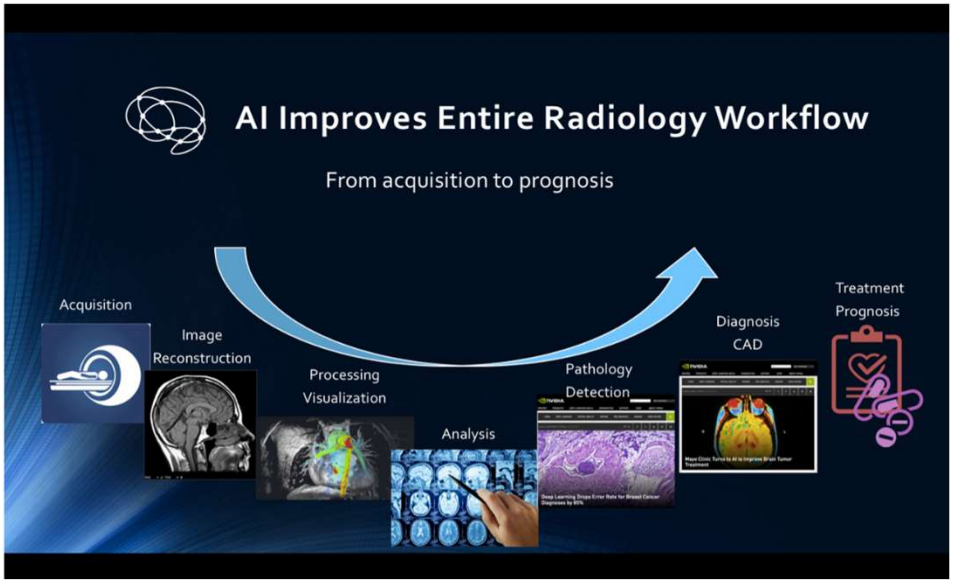


https://www.arterys.com/

THE LINUX FOUNDATION

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

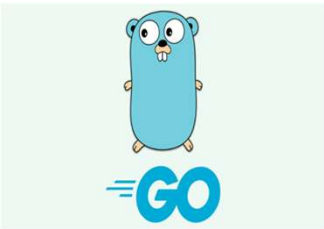


https://subtlemedical.com/ai-is-starting-to-change-radiology-for-real/

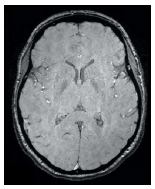
Project Objective

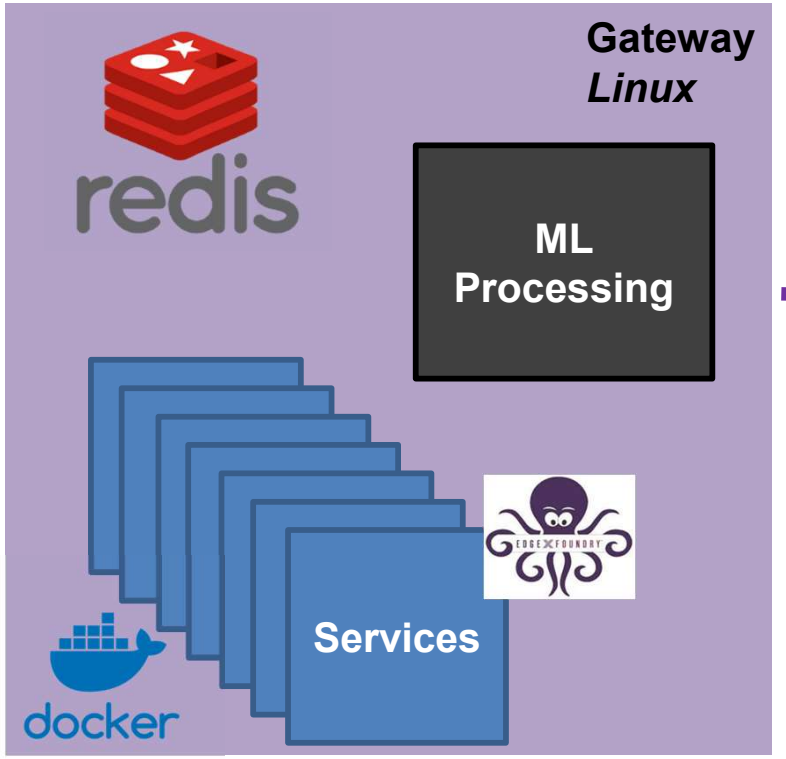
Project Objective

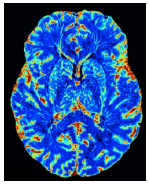
- Automated image transfer, processing and comparison
- Go Microservices & Containerized
- Reference Implementation (RI) using Open-Source Software
- AI-Assisted, processing pipeline reducing barrier at the Edge



Health Care Solution using AI/ML







THE LINUX FOUNDATION

Architecture

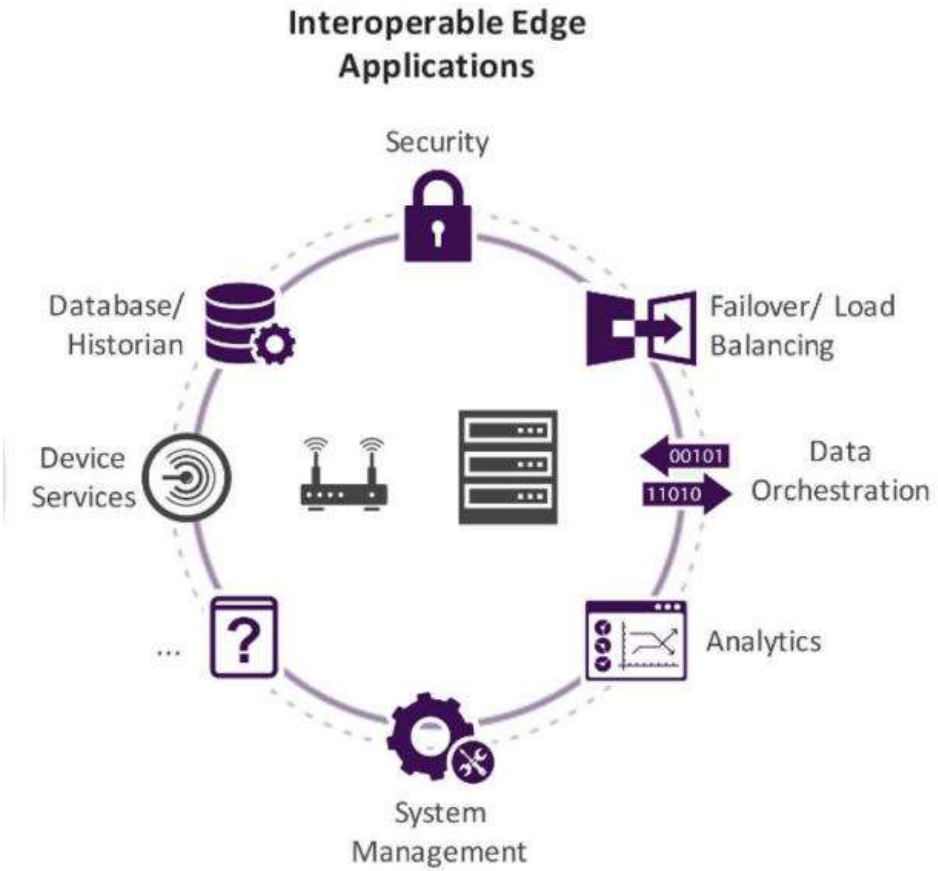
Terminology

- 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:

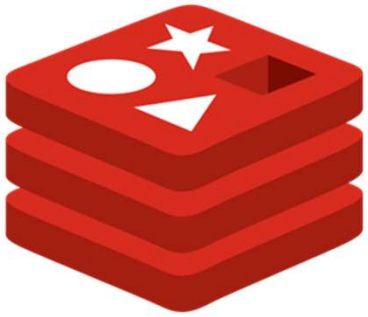
- Hosted by LF Edge
- Framework for IOT edge compute
- Application Service API



Storage

Redis:

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



redis

ML Pipeline Management

TIBCO Project Air:

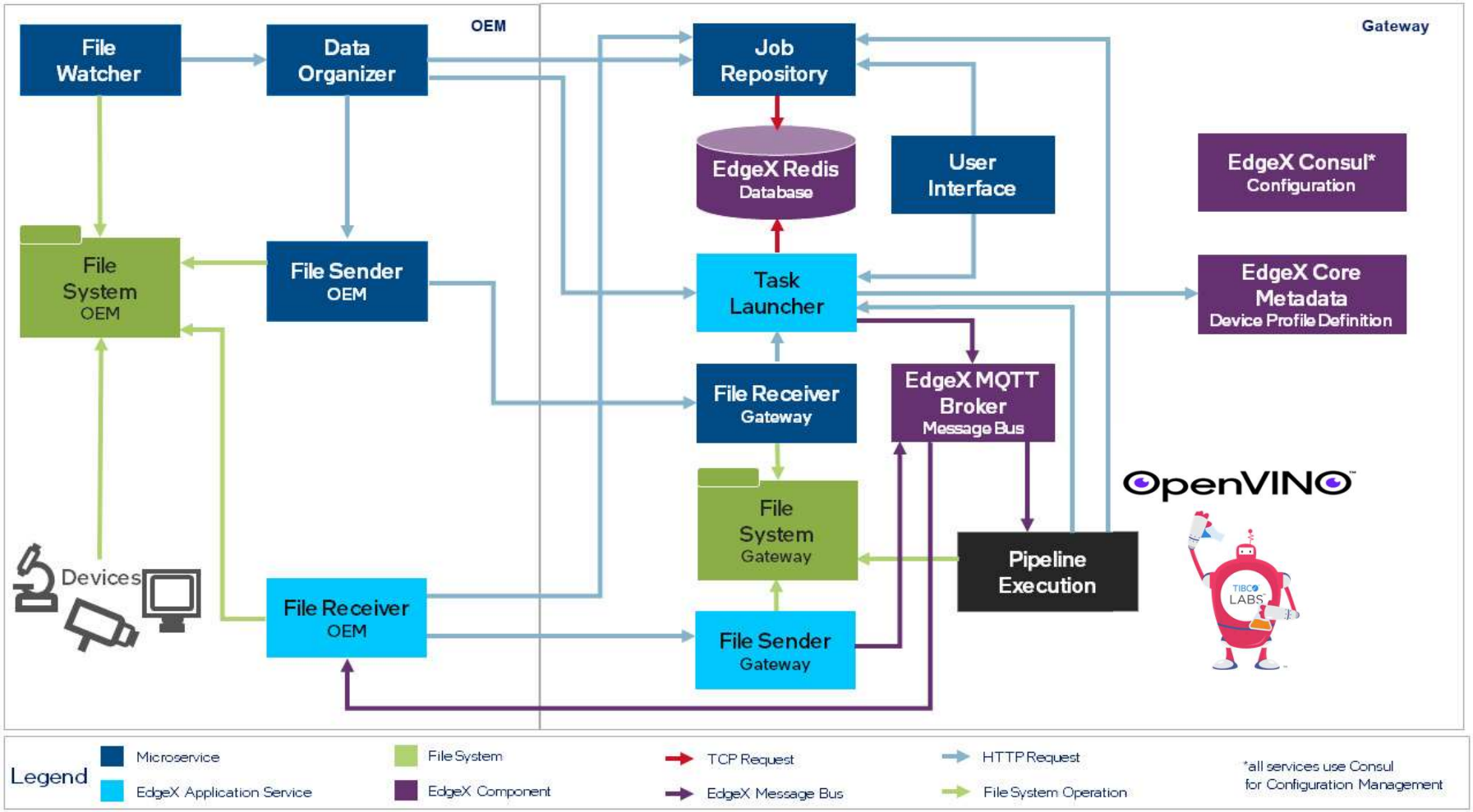
- Build, Configure & Modify Pipelines
- Uses EdgeX Foundry

Intel® Distribution of OpenVINO™ Toolkit

- Open Source Tool for optimized Inferencing on Intel

OpenVINO™

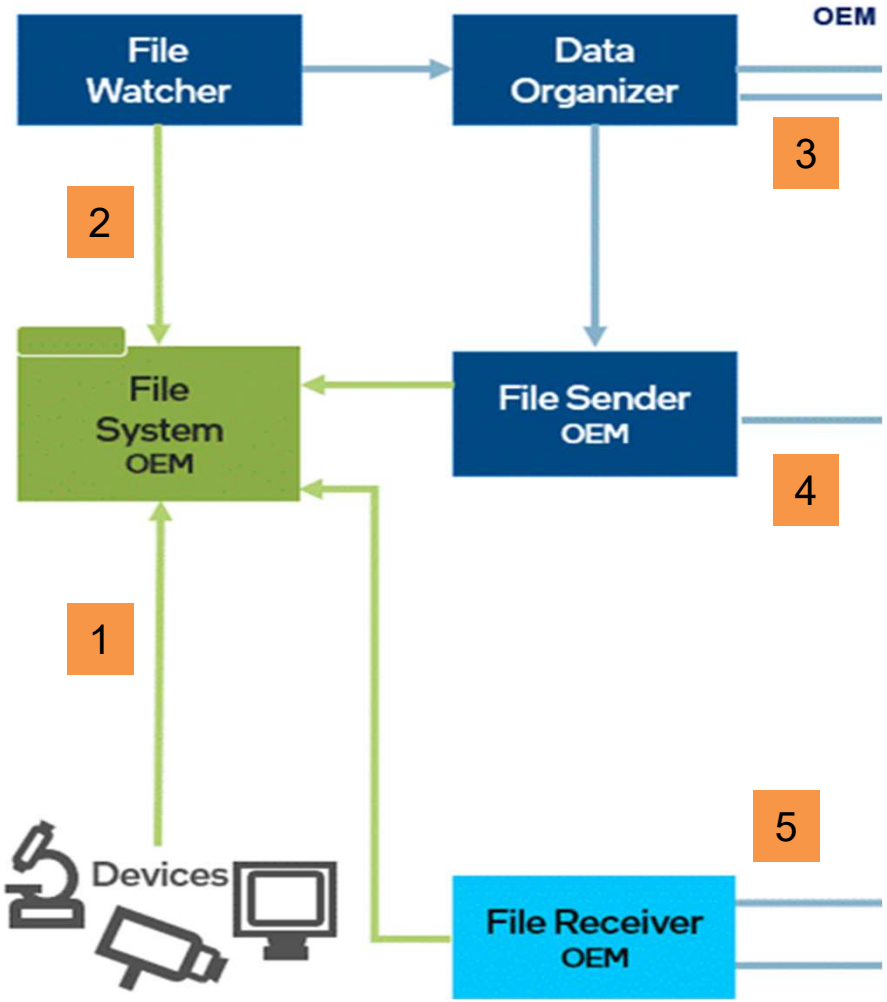
Architecture



OEM Microservices

Watch specified folders for images

Send REST call to Data Organizer with Job



Routes calls to Job repository

Sends jobs & files received to Gateway

Subscribes & pulls output files

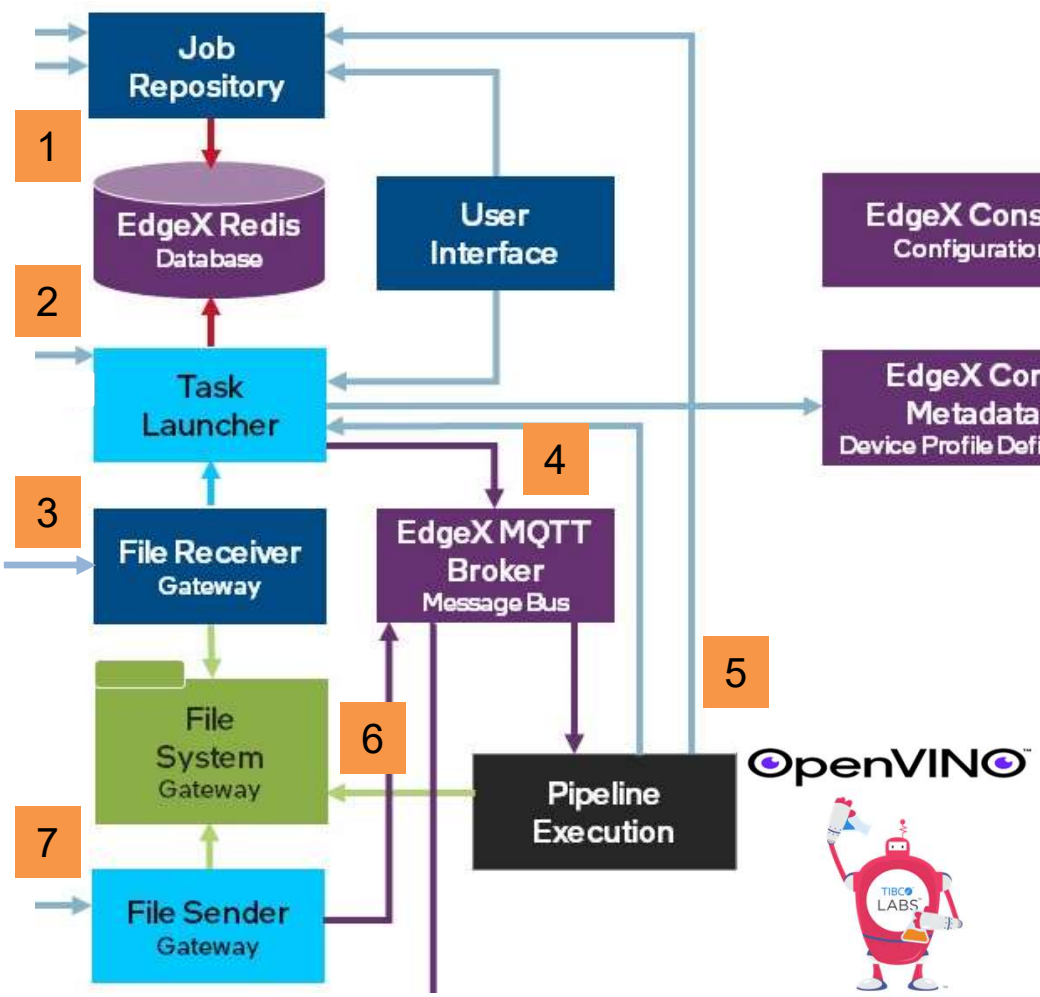
GW Microservices

Create Job in Redis

Verifies matching task & send to pipeline

Receiver job & file & write to file system

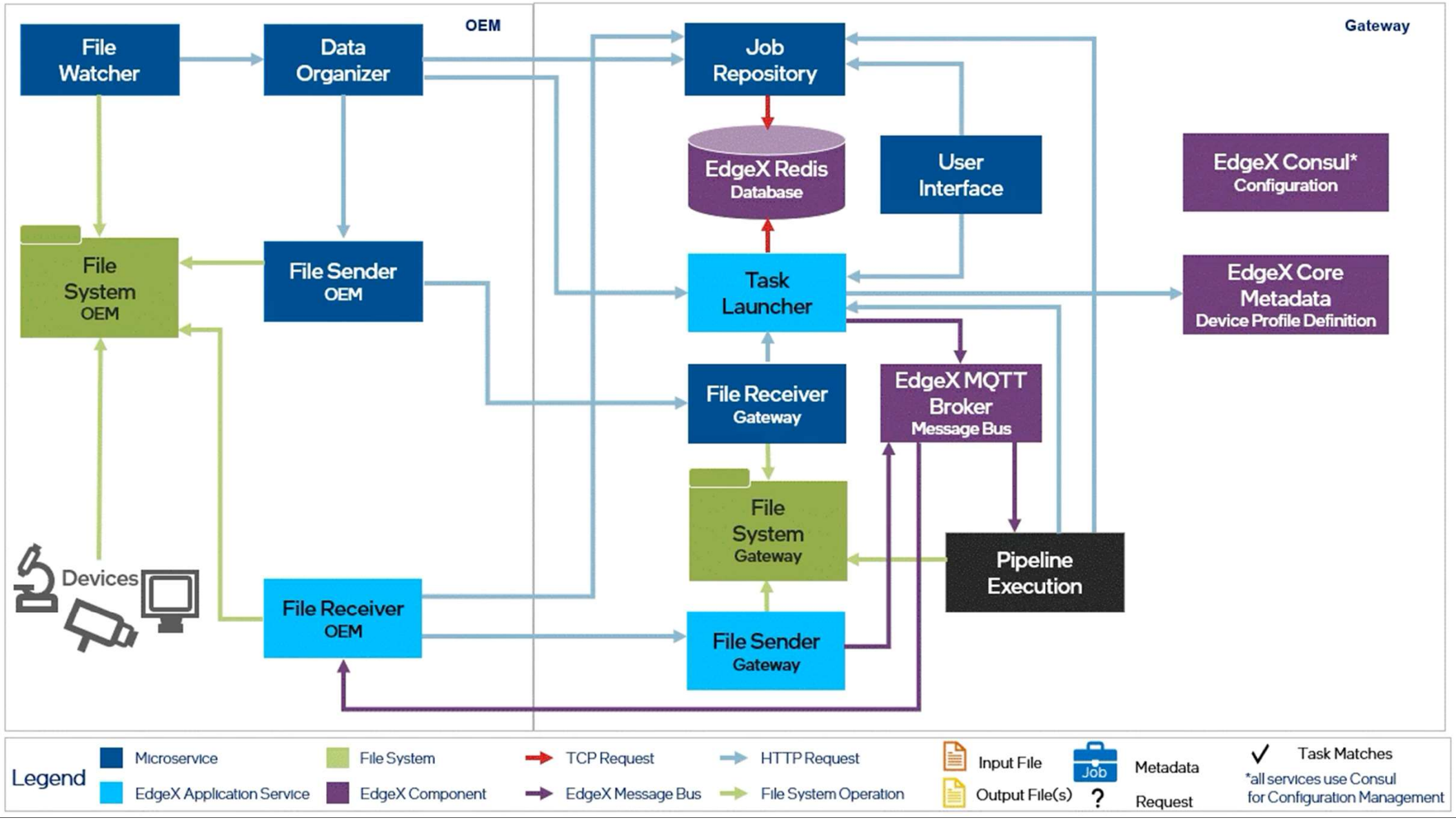
Publish job & accept request for output files



Service discovery

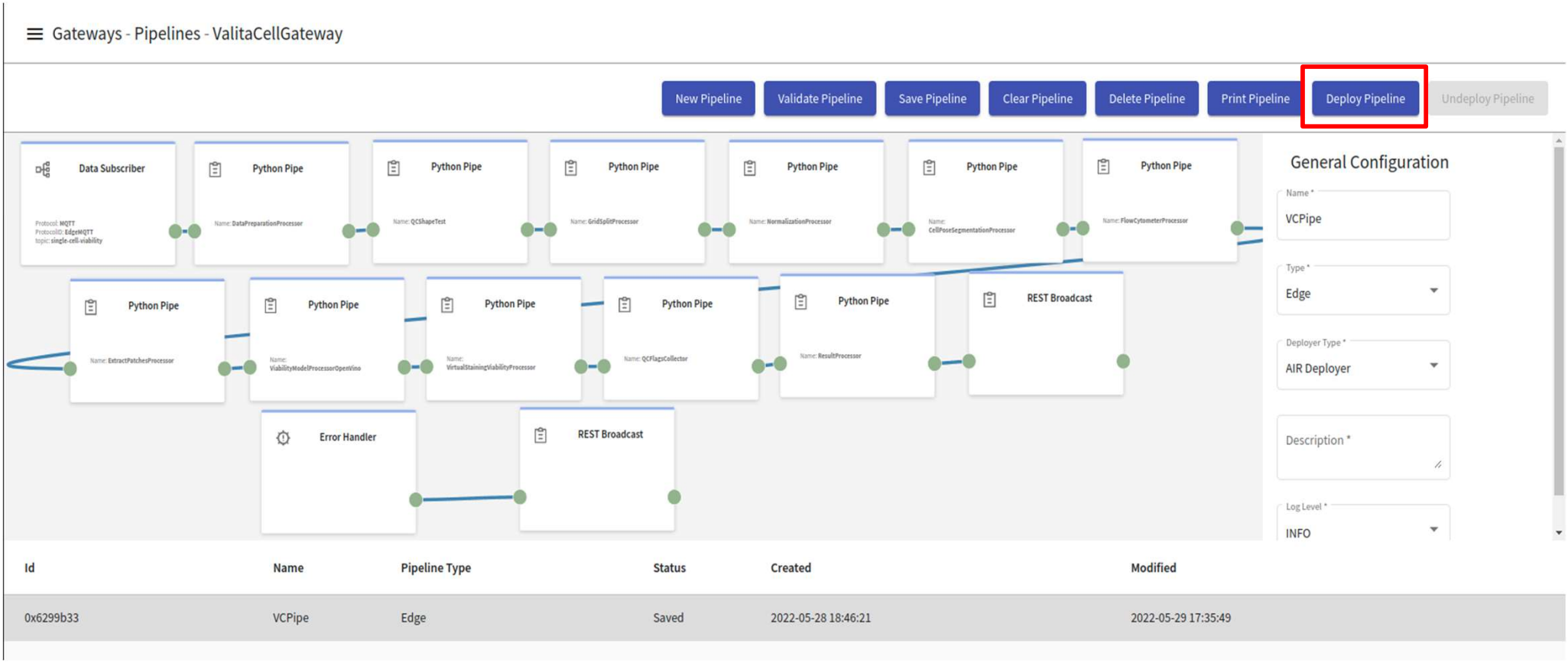
Service metadata

Execute custom models to images


OEM
Gateway
File Watcher
Data Organizer
Job Repository
EdgeX Redis Database
User Interface
EdgeX Consul* Configuration
File System OEM
File Sender OEM
Task Launcher
EdgeX Core Metadata Device Profile Definition
File Receiver Gateway
EdgeX MQTT Broker Message Bus
File System Gateway
Pipeline Execution
Devices
File Receiver OEM
File Sender Gateway
Legend
Microservice
EdgeX Application Service
File System
EdgeX Component
TCP Request
EdgeX Message Bus
HTTP Request
File System Operation
Input File
Output File(s)
Job
Metadata
Request
Task Matches
*all services use Consul for Configuration Management

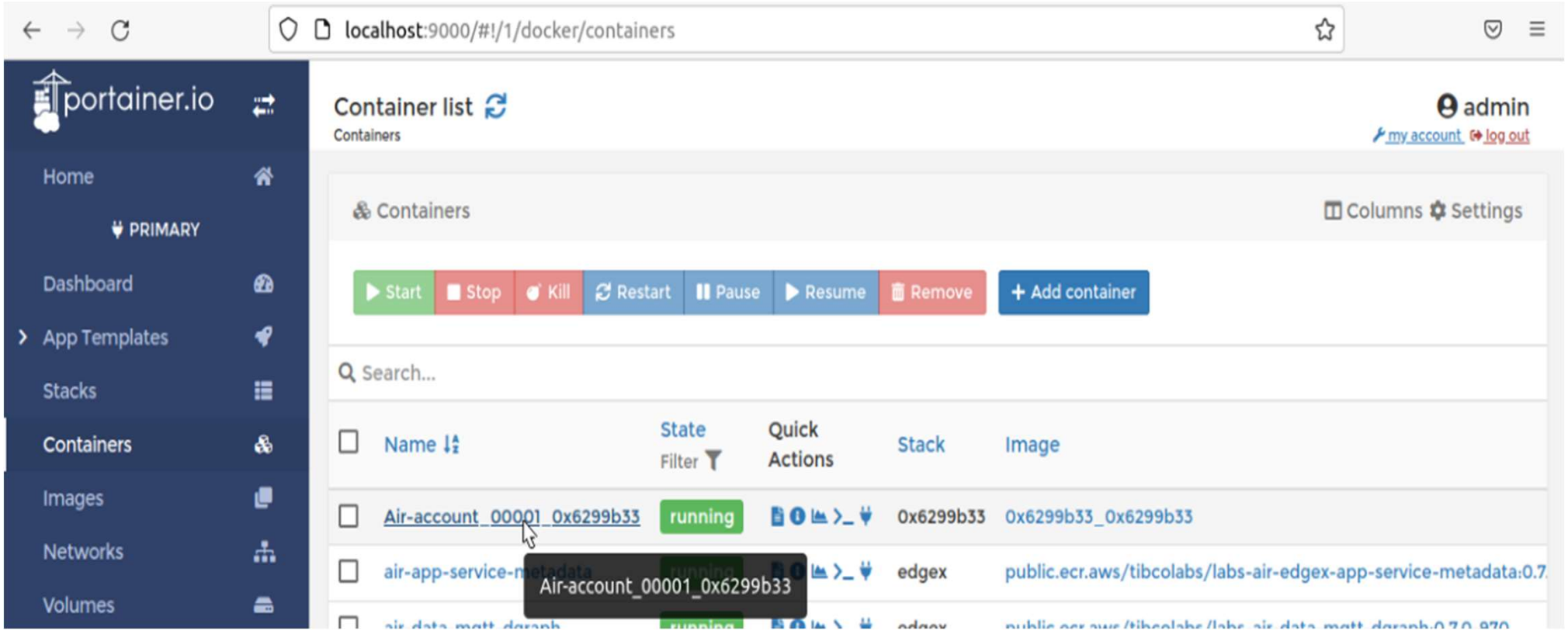
UI & Features

Project Air UI



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

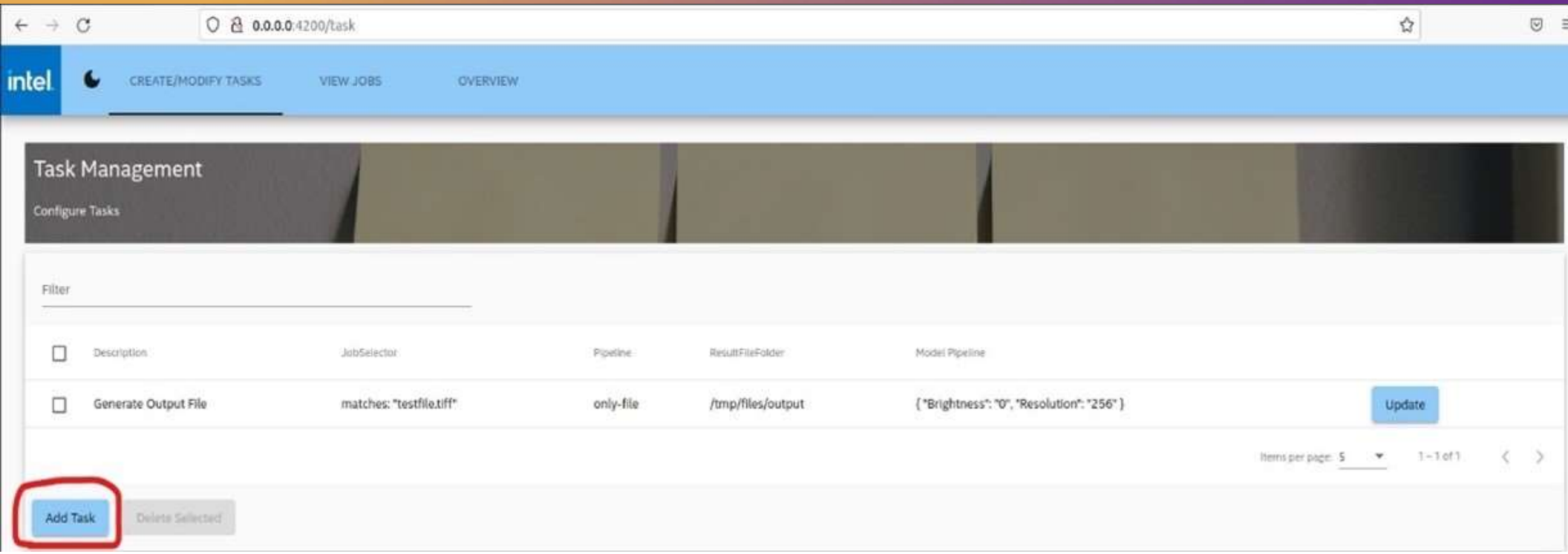


User Interface

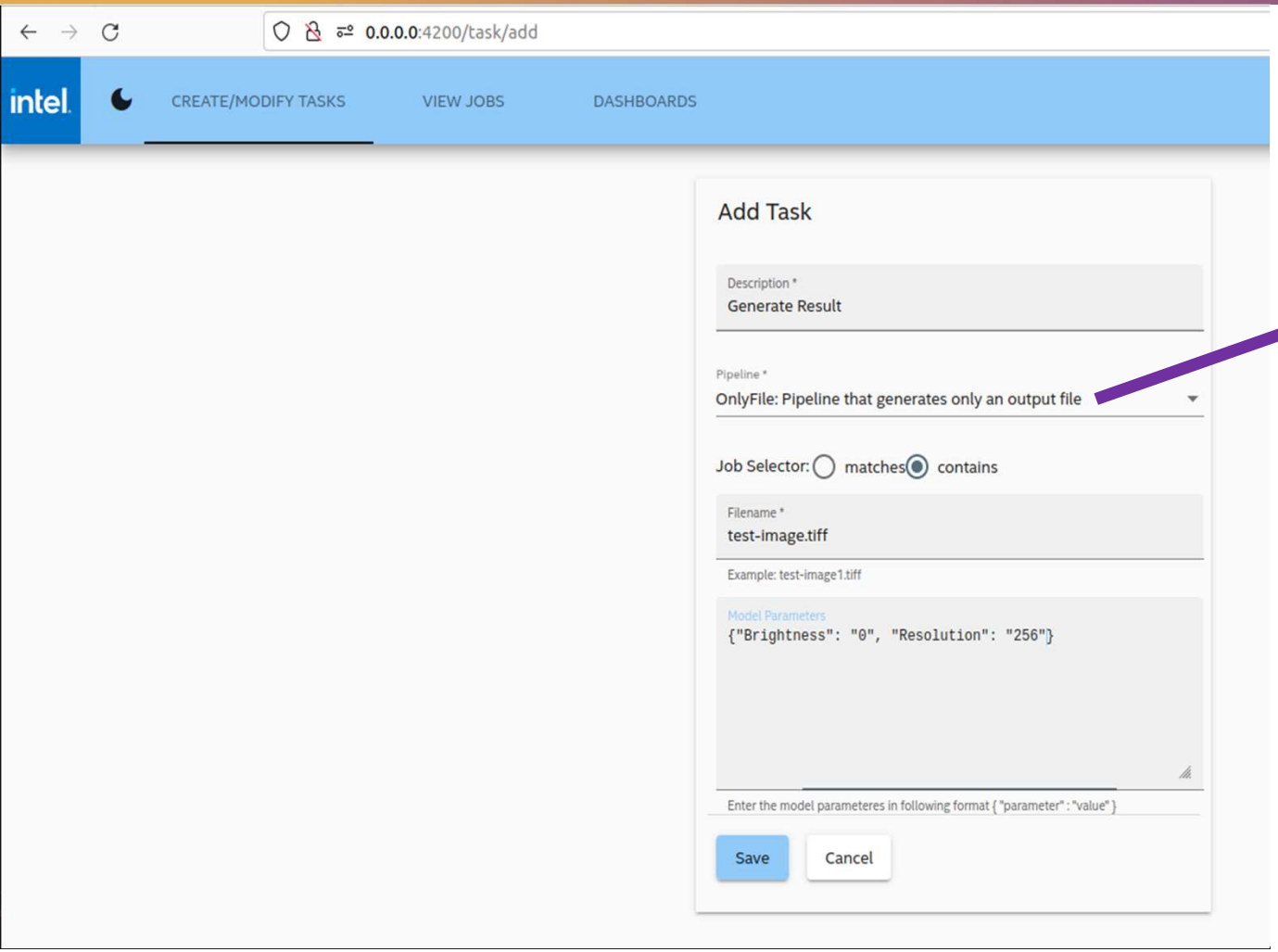
Angular UI:

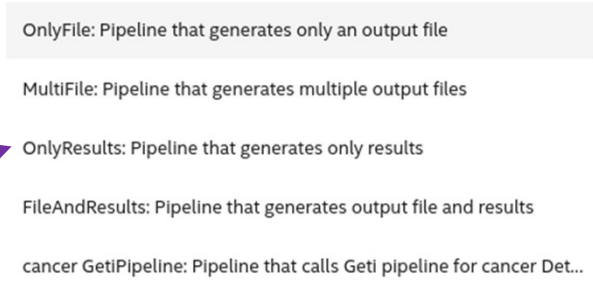
- Task Creation
- Job Monitoring
- Features:
  - Observability
  - Internationalization

Create Task

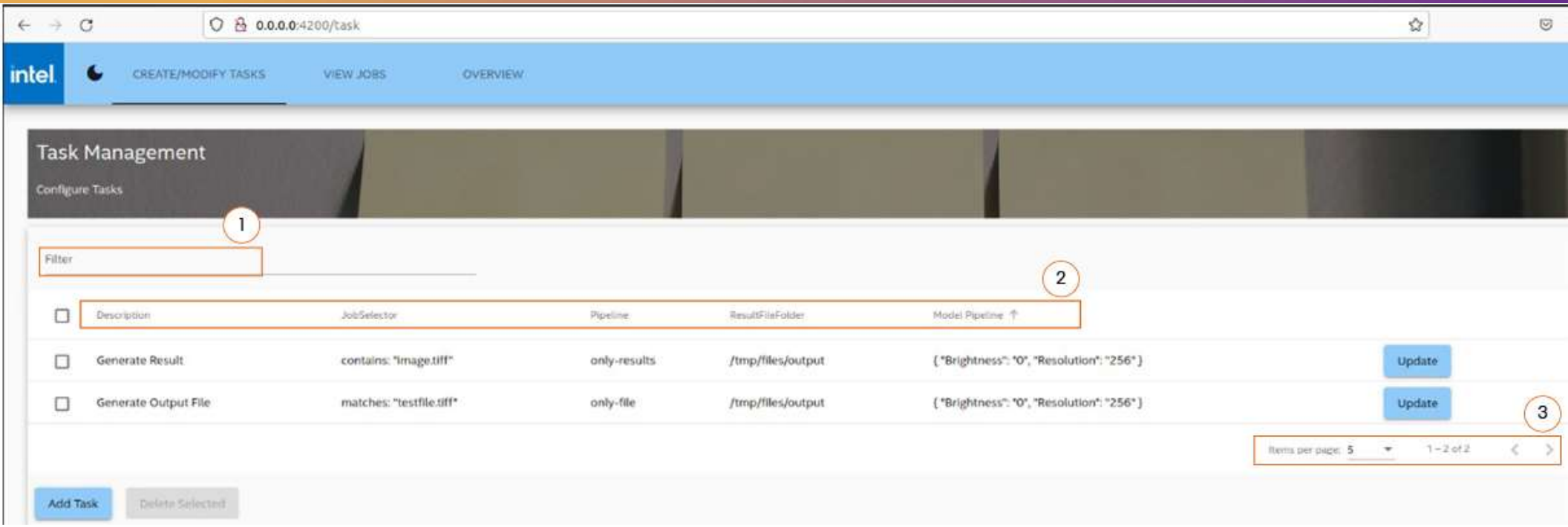

0.0.0.0:4200/task
intel
CREATE/MODIFY TASKS
VIEW JOBS
OVERVIEW
Task Management
Configure Tasks
Filter
Description
JobSelector
Pipeline
ResultFileFolder
Model Pipeline
Generate Output File
matches: "testfile.tiff"
only-file
/tmp/files/output
{ "Brightness": "0", "Resolution": "256" }
Update
Items per page: 5
1 – 1 of 1
Add Task
Delete Selected

Select Pipeline

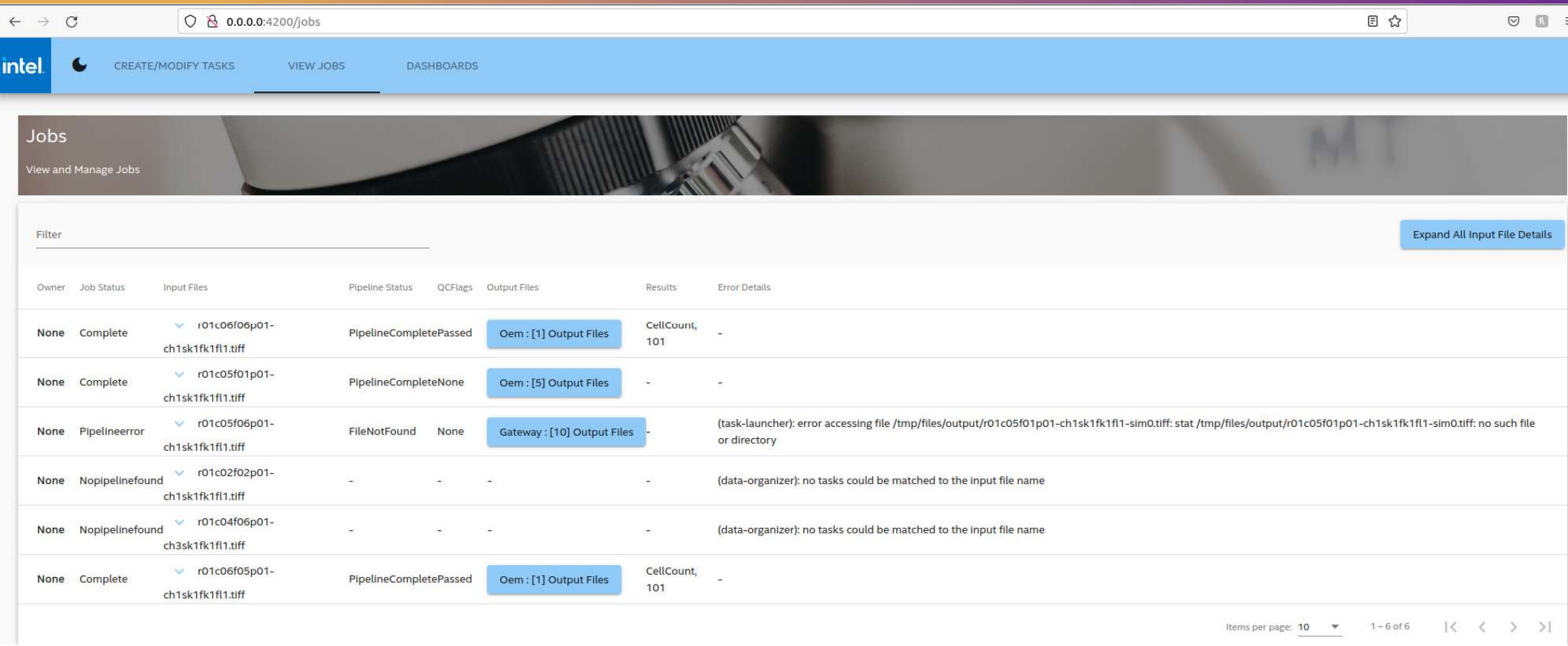




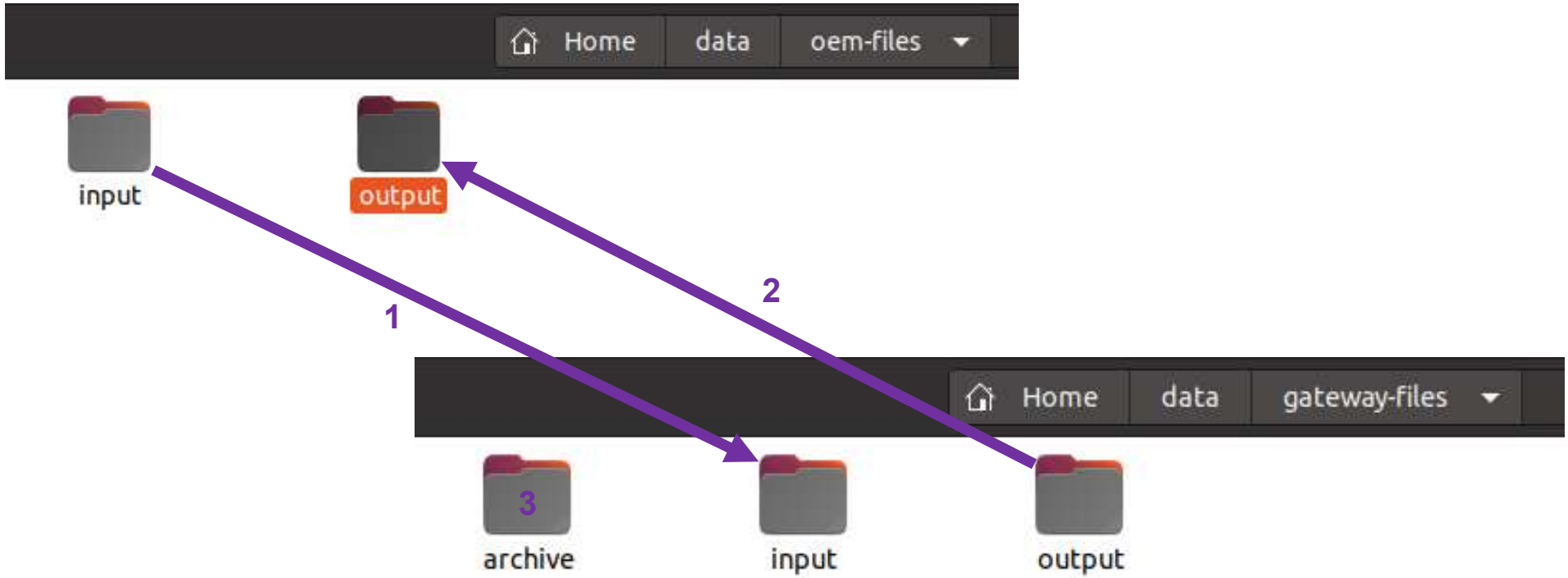
Display Task



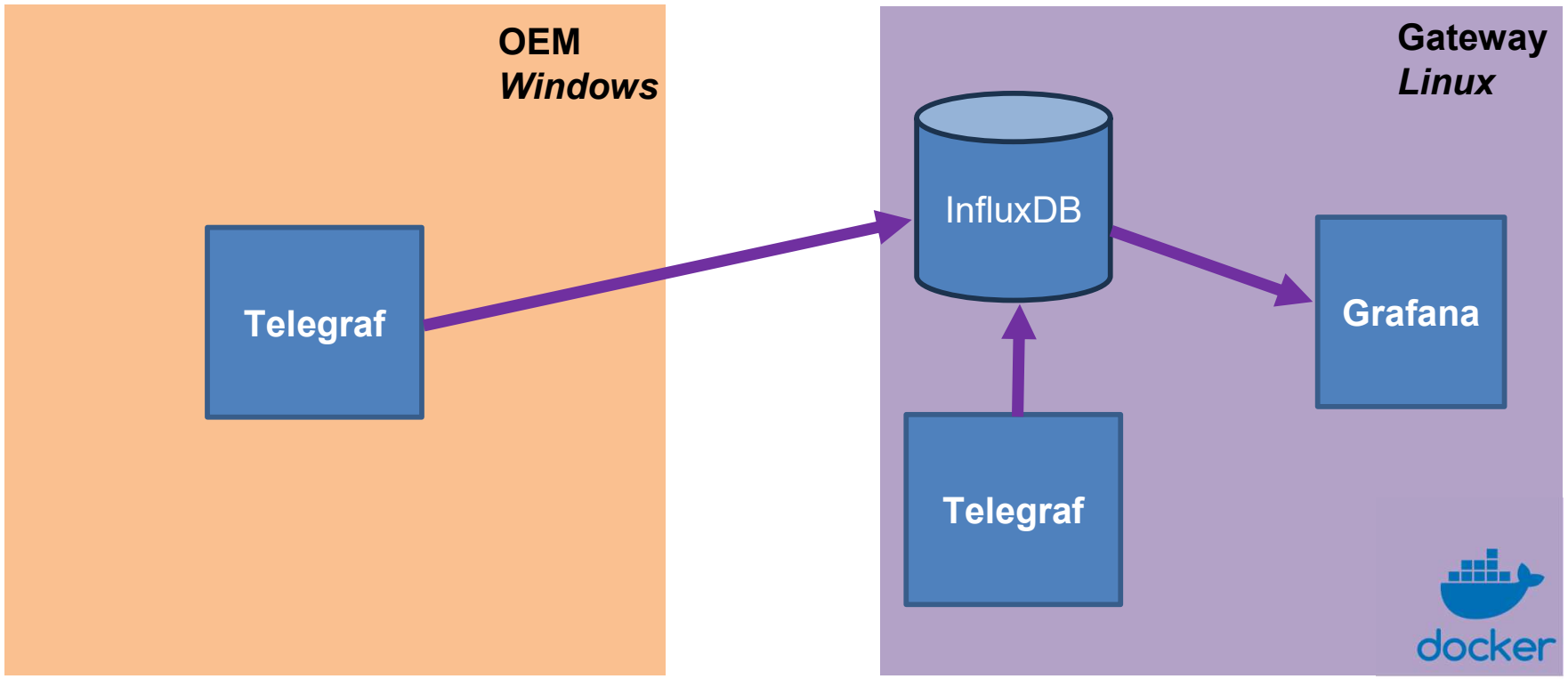
Display Jobs



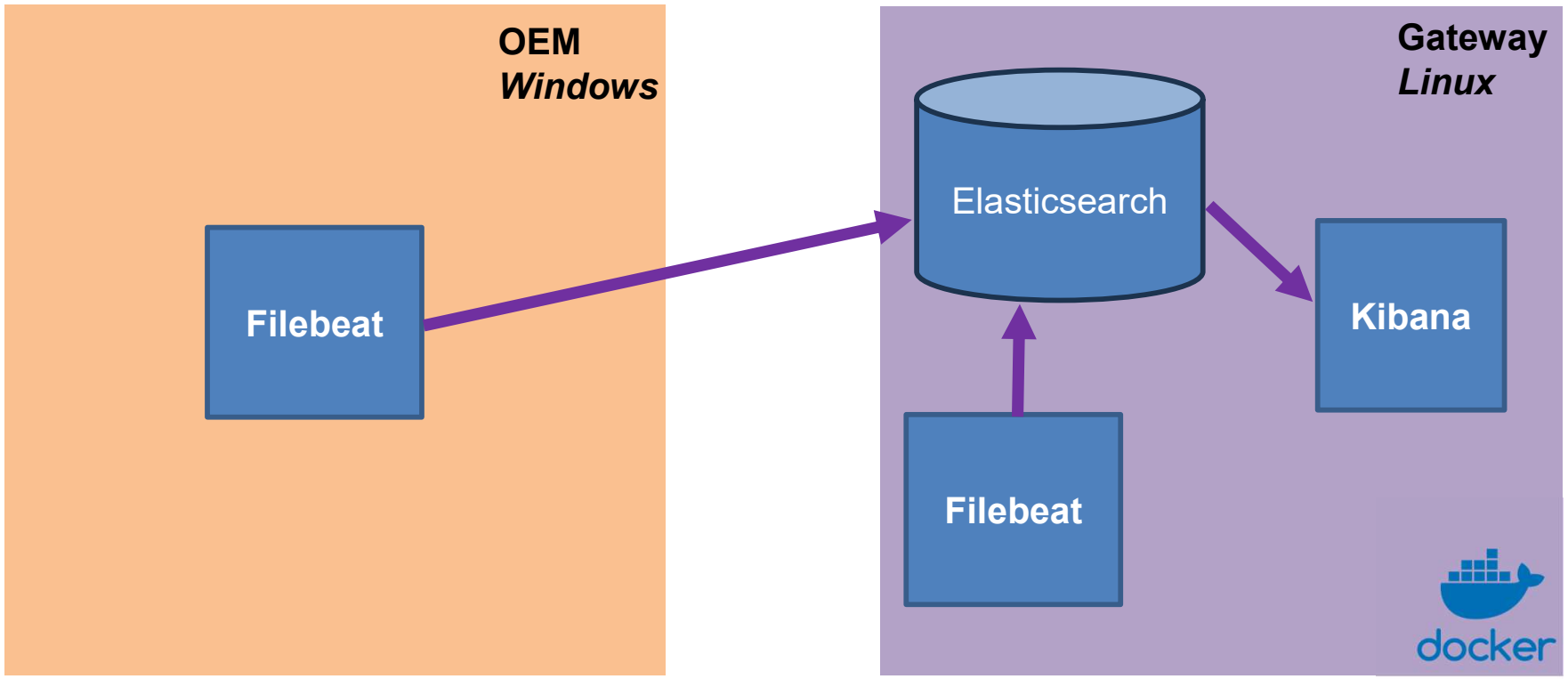
Input & Output File Folders – OEM & Gateway



Observability

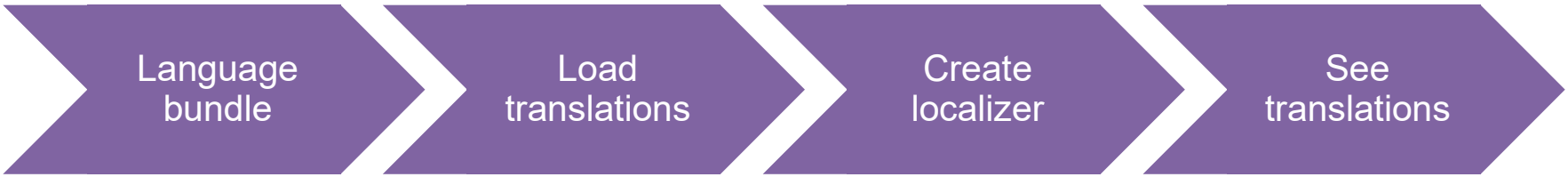


Log Analytics

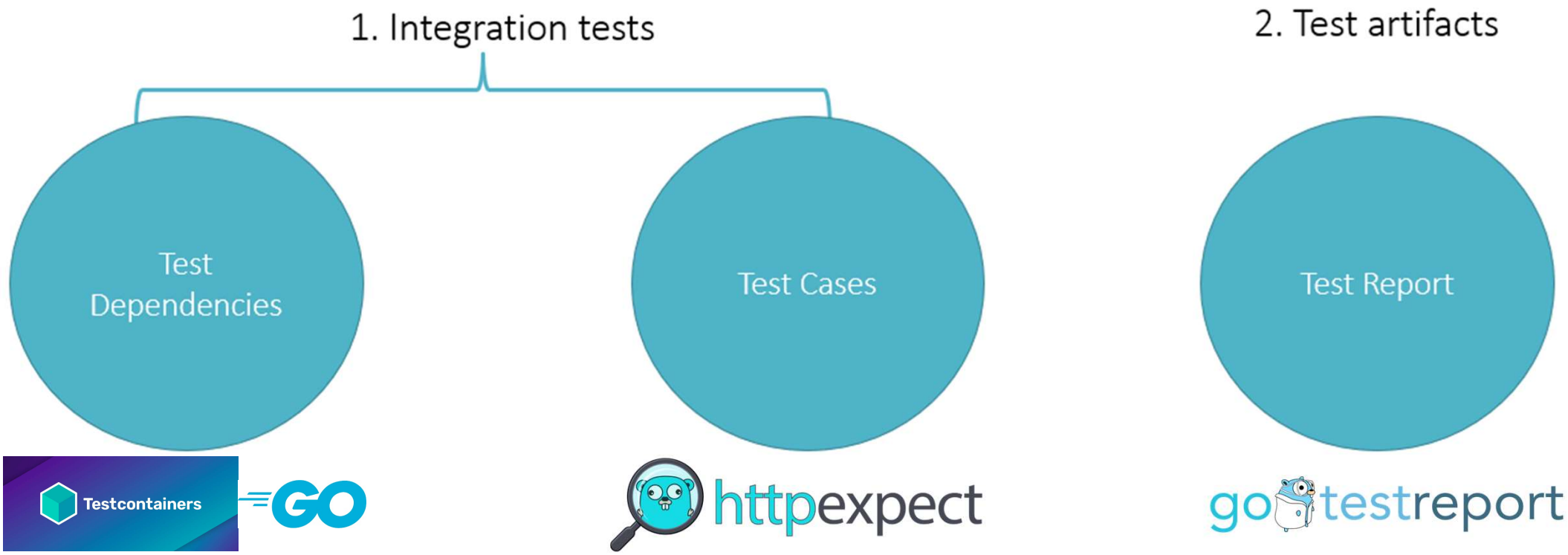


Internationalization

- Localization of many Job fields to Chinese
- Go-i18n package
- Update headers:
  - Accept-Language: en-zh



Integration Testing – Open Source Go Packages



THE LINUX FOUNDATION

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:

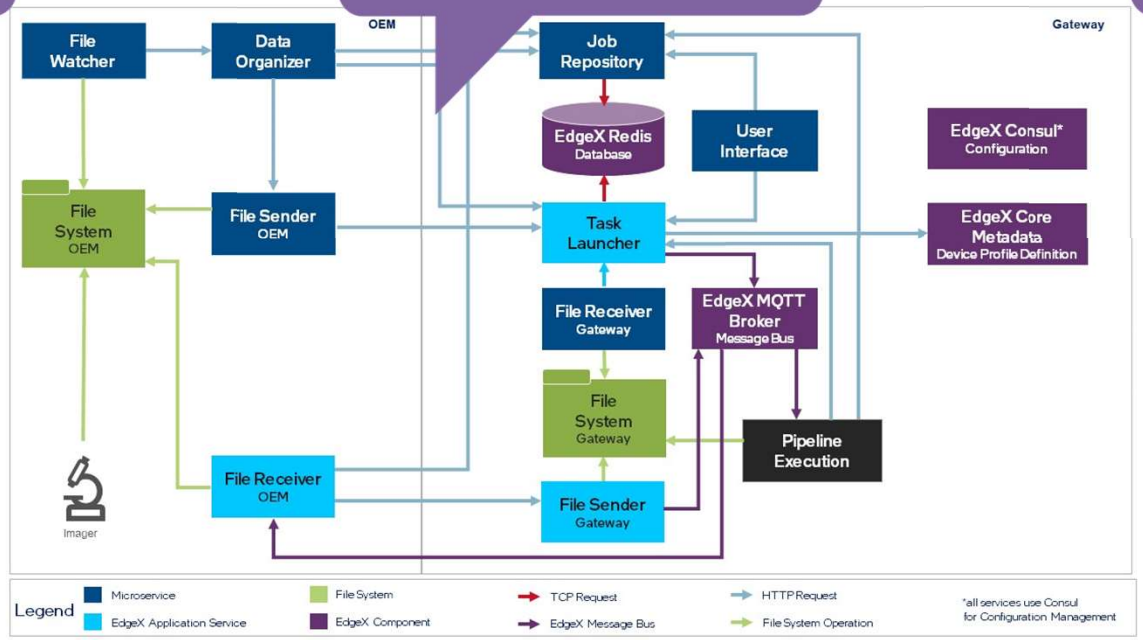
- Gateway services:

Service Wait Strategy

Deployment time

What ifs?

Integration test time



THE LINUX FOUNDATION

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

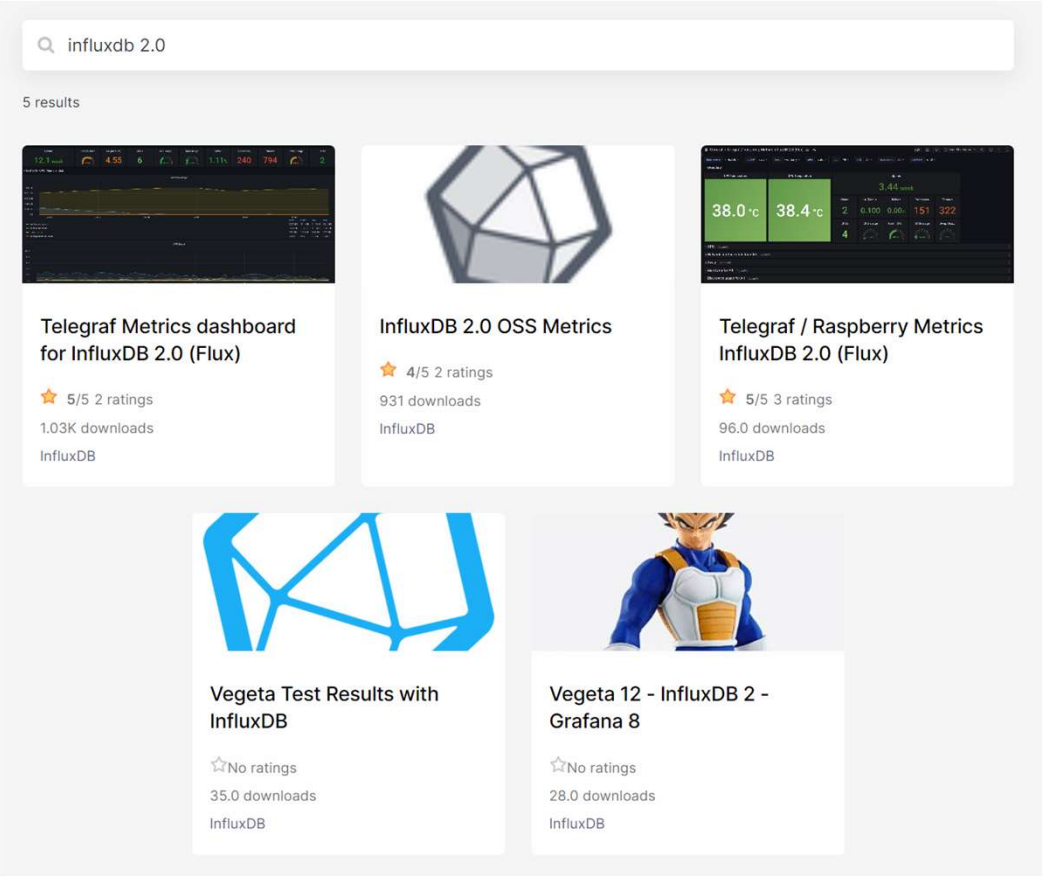
4 seconds

Observability

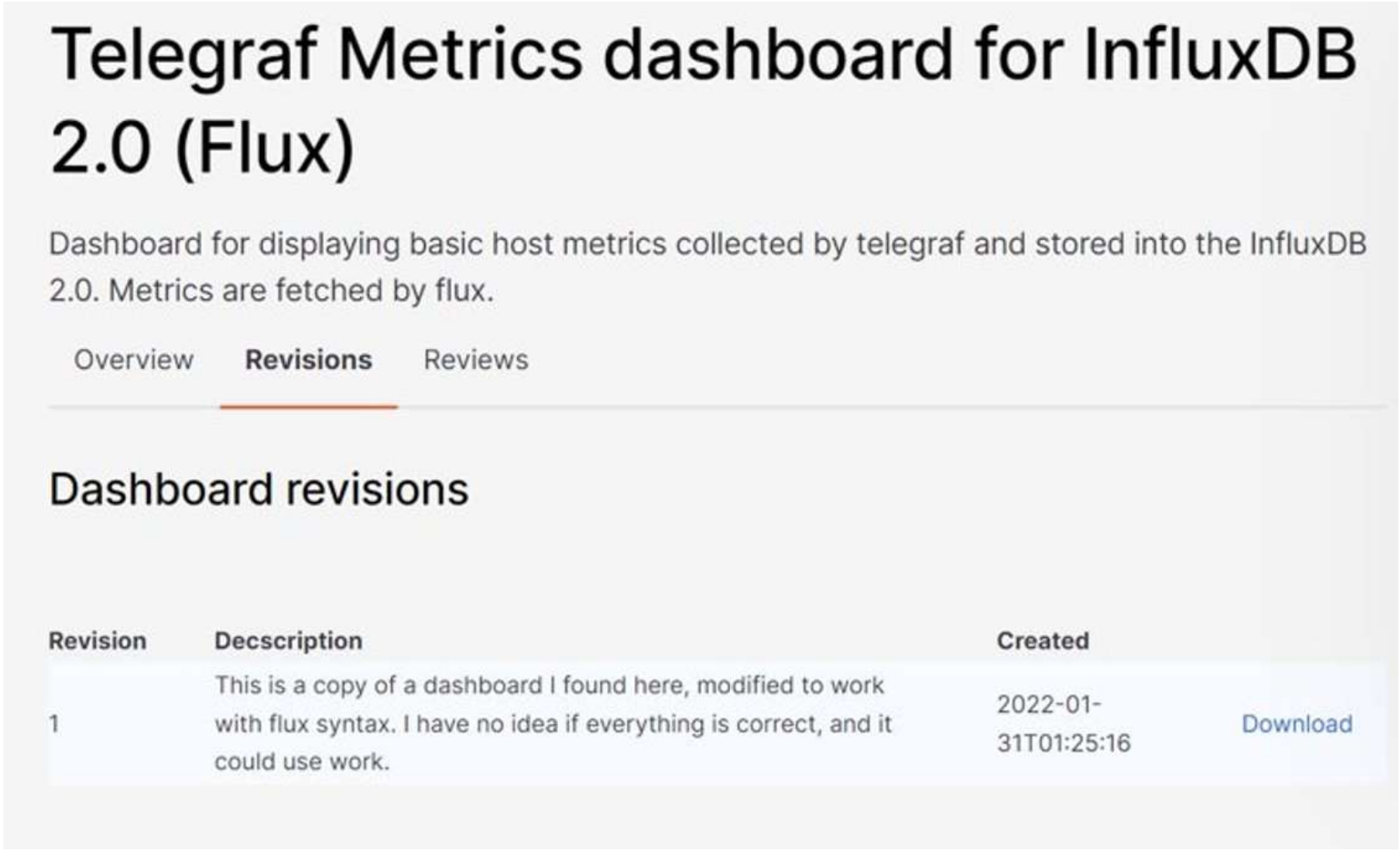
InfluxDB 2.0

Flux-styled queries

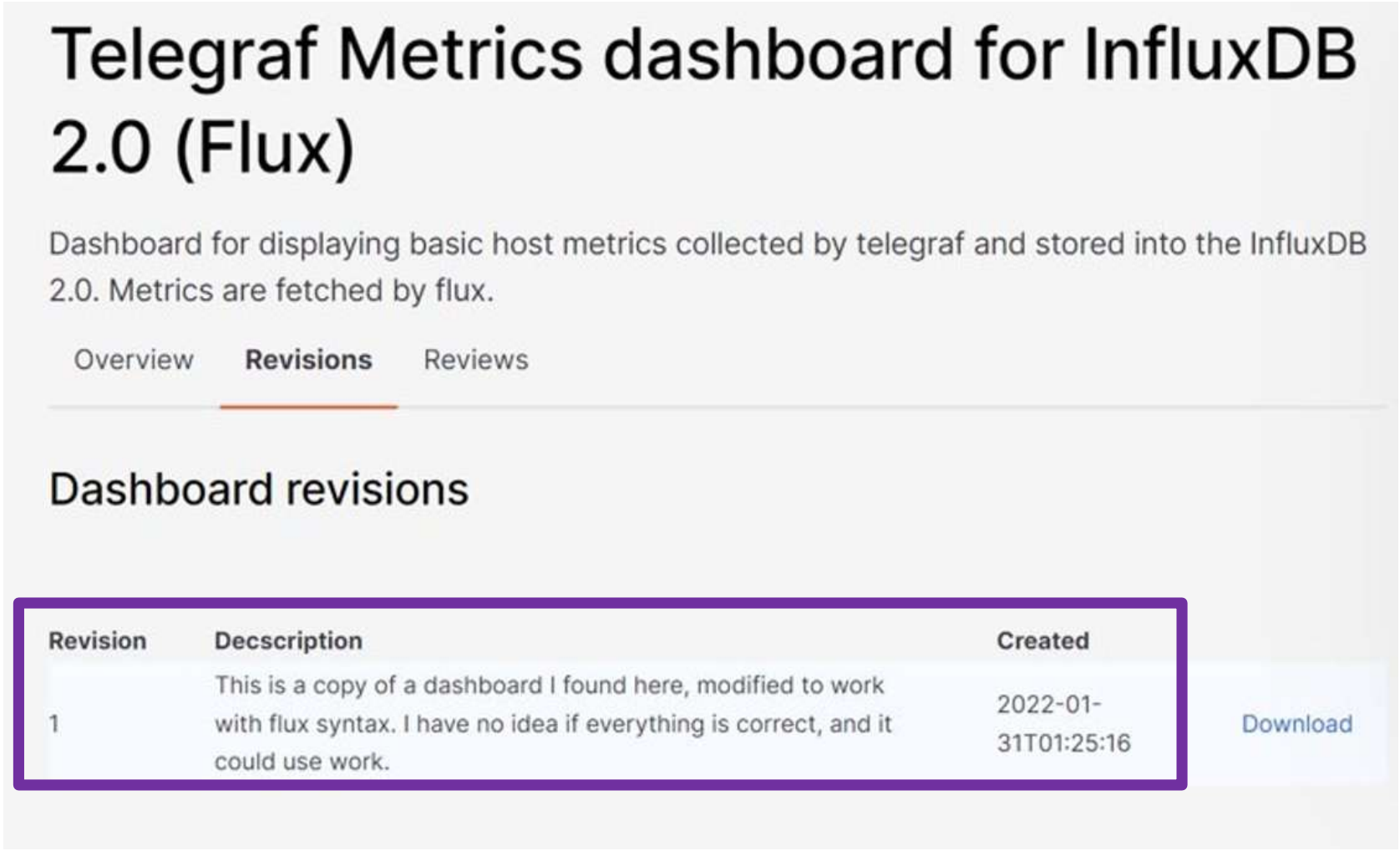
Cutting edge



Observability

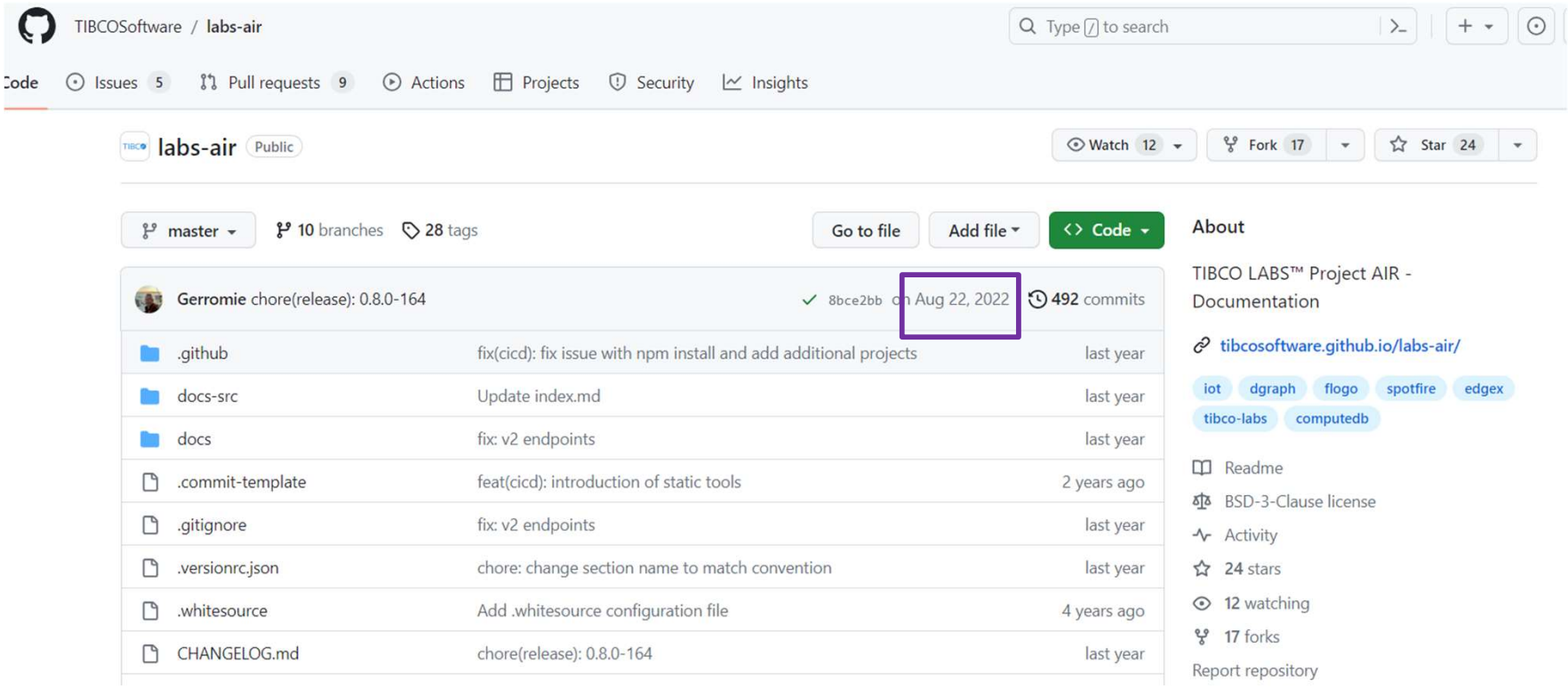


Observability



Open Source Project Shifts

TIBCO Project Air



Conclusion

Conclusion

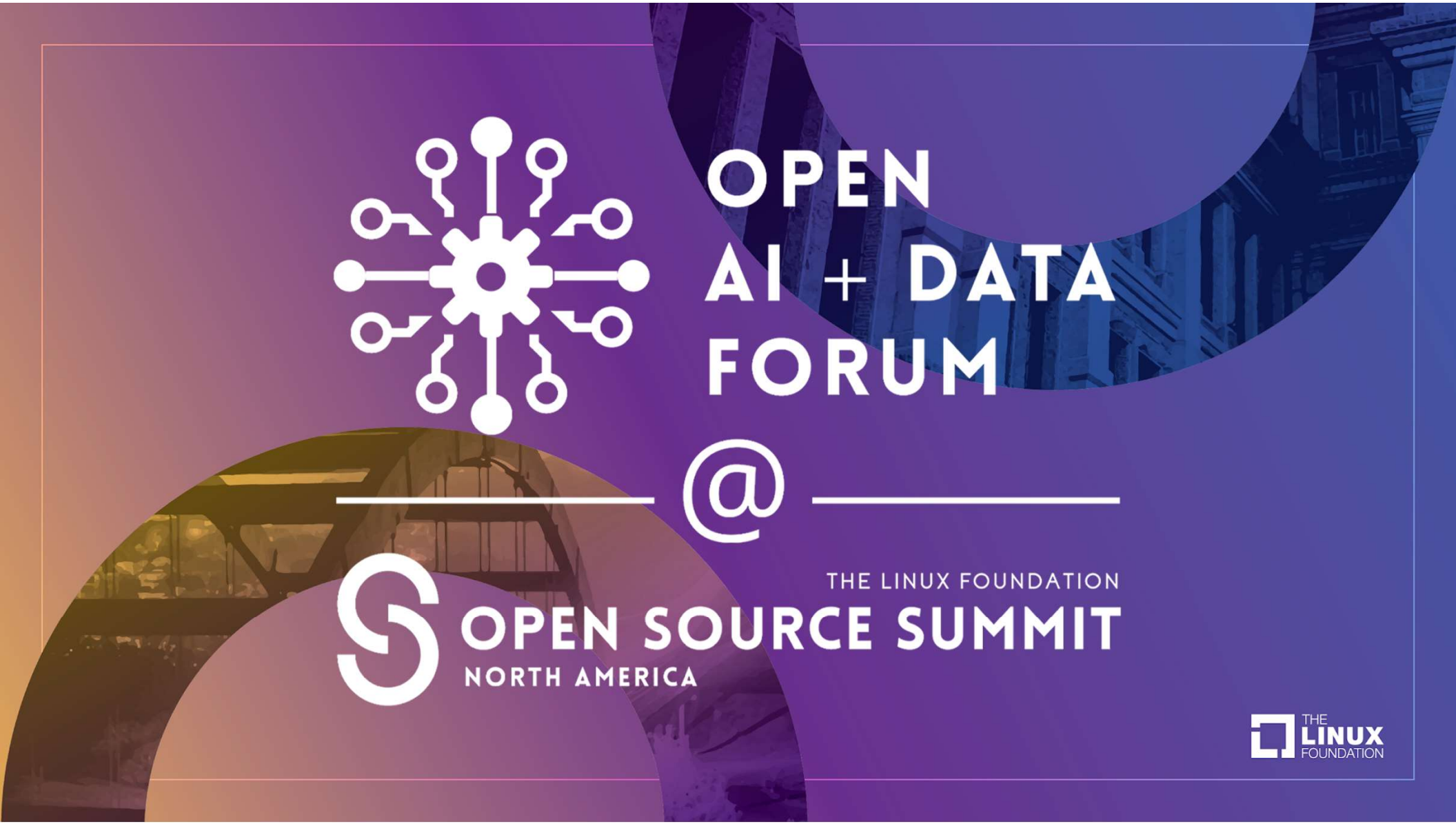
- Replaced Project Air with Open-Source Tool - BentoML
- Released as Open-Source Sample in July
- AiCSD – AI Connect for Scientific Data - https://github.com/intel/AiCSD
- Successfully Deployed customer's lab in Colorado
- Recorded Webinar on using OpenVINO optimization to unlock AI performance for CellAI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

Thank You

Questions?

Neethu Elizabeth Simon
Senior Software Engineer
Intel Corporation

THE LINUX FOUNDATION


OPEN
AI + DATA
FORUM
@
THE LINUX FOUNDATION
OPEN SOURCE SUMMIT
NORTH AMERICA
THE LINUX FOUNDATION