Alibaba Cloud | Worldwide Cloud Services Partner

AI made easy with Apache Flink + AI Flow

Jiangjie (Becket) Qin

Staff Software Engineer & Senior Manager

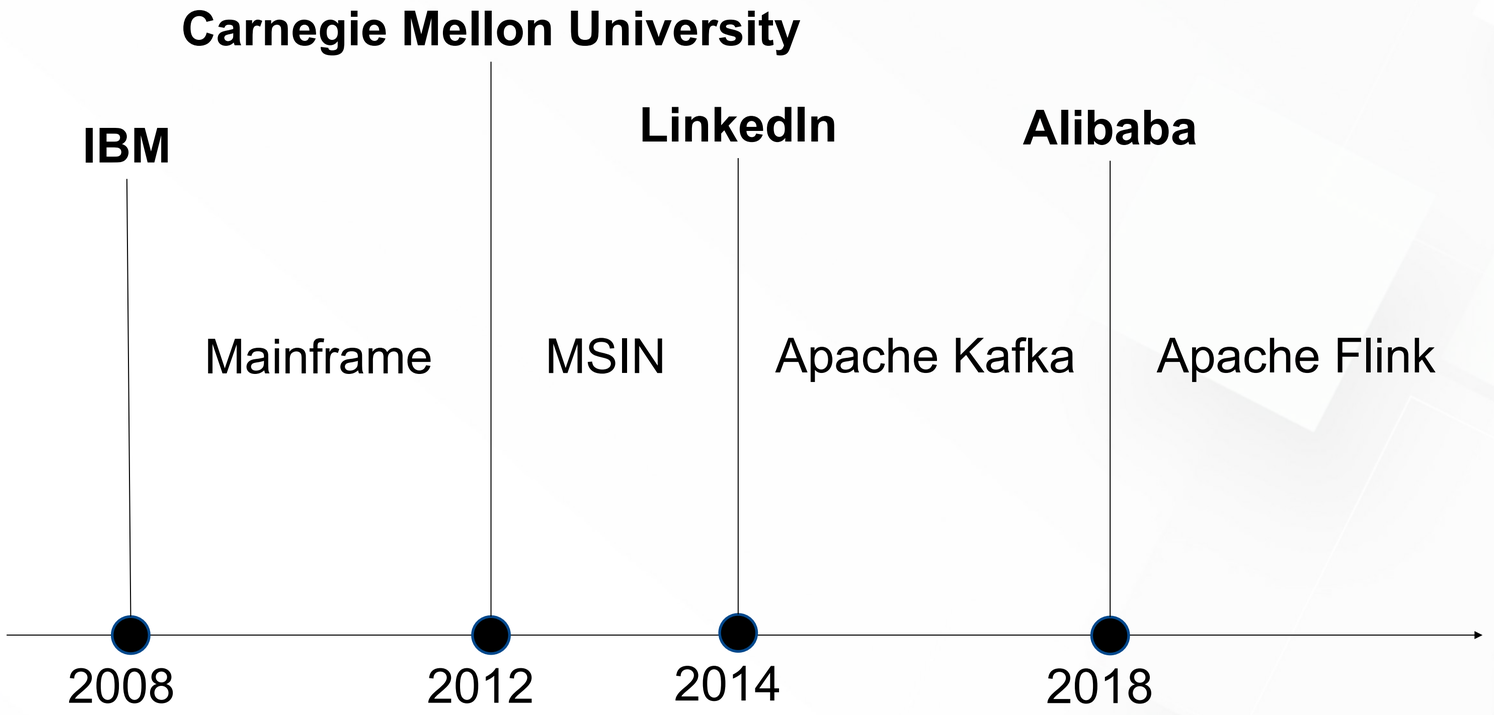
24/04/2020

About Me

Alibaba Cloud | Worldwide Cloud Services Partner

Alibaba Inc.
Staff Software Engineer
Senior Manager

Apache Flink PMC Member
Apache Kafka PMC Member



Agenda

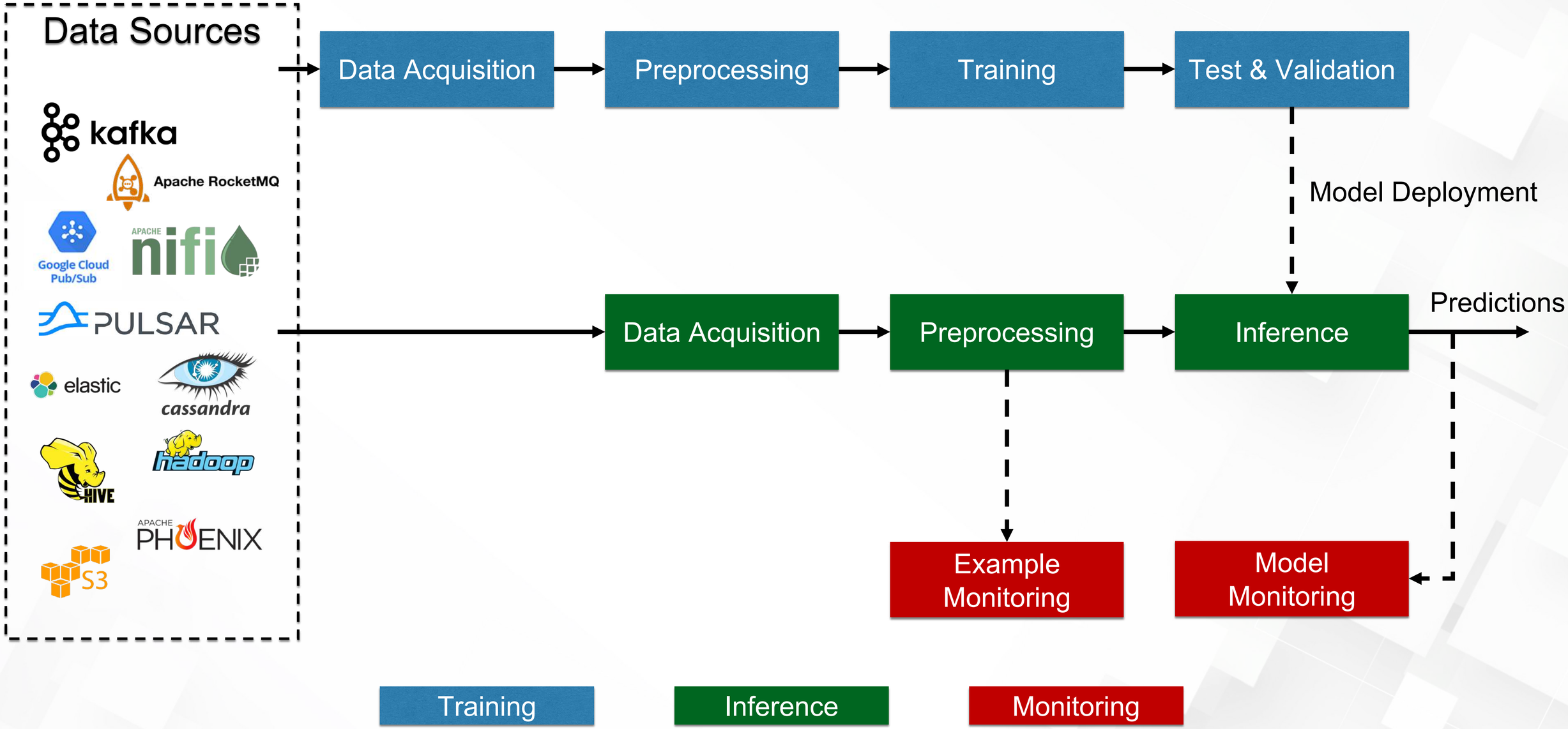
- What is AI Flow
- AI Flow Introduction
- Flink AI Flow
- Summary

What is AI Flow

- An overview of AI workflow
- Challenges in AI workflow
- Meet AI Flow

An Overview of AI Workflow

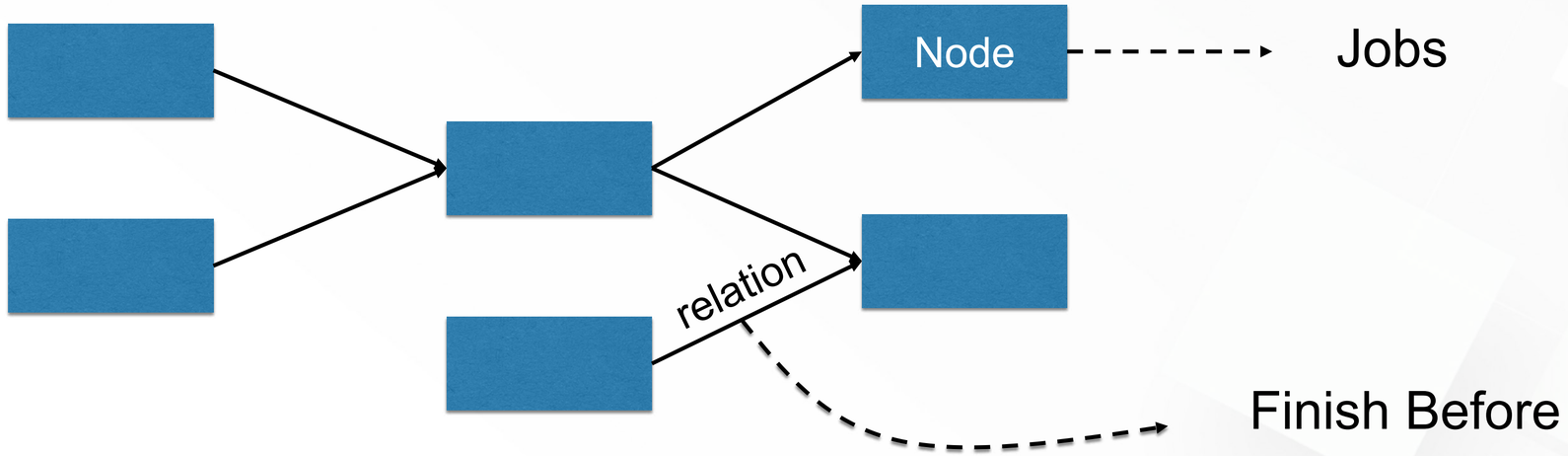
Alibaba Cloud | Worldwide Cloud Services Partner



Understand a Workflow

Alibaba Cloud | Worldwide Cloud Services Partner

Workflow Definition

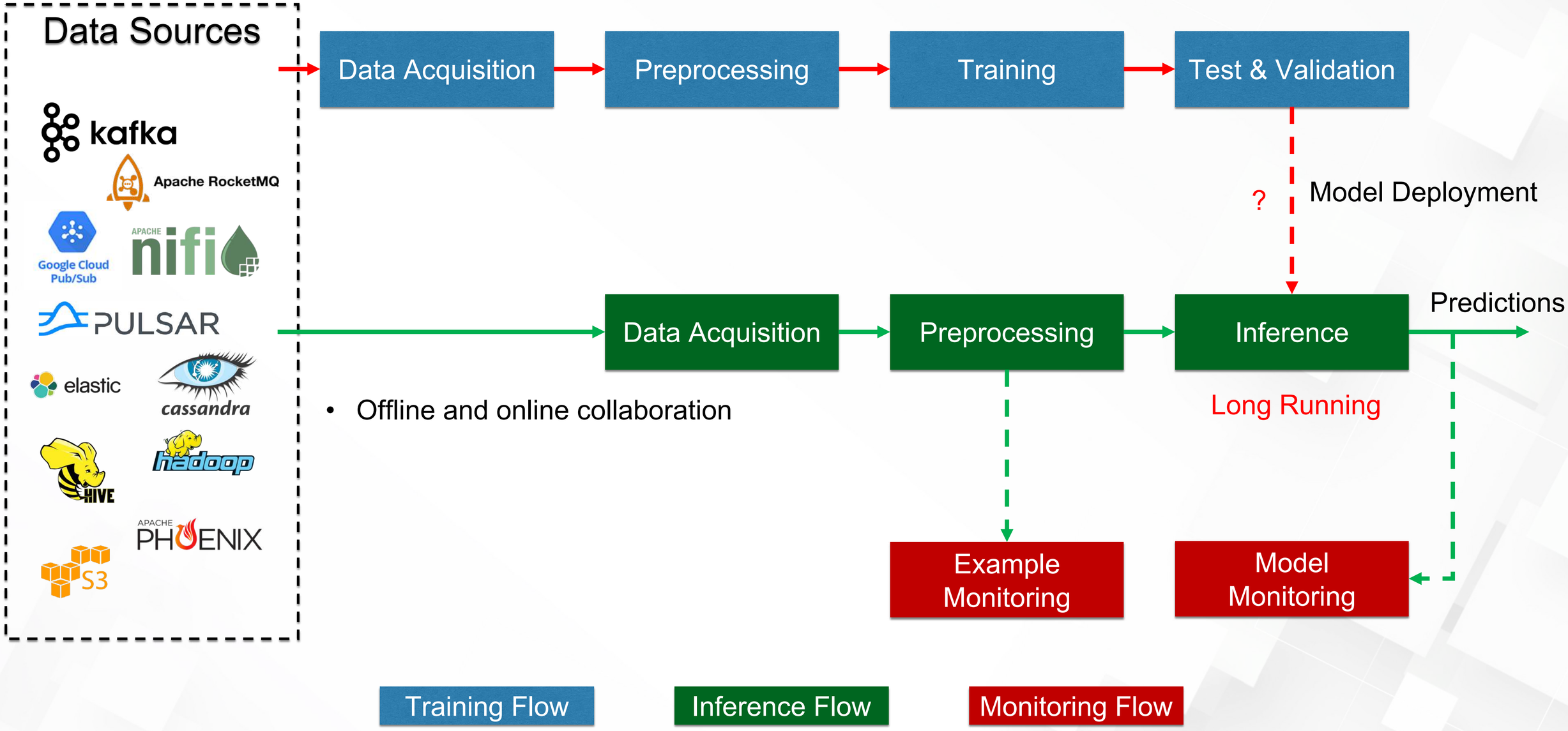


Workflow Execution

Scheduler

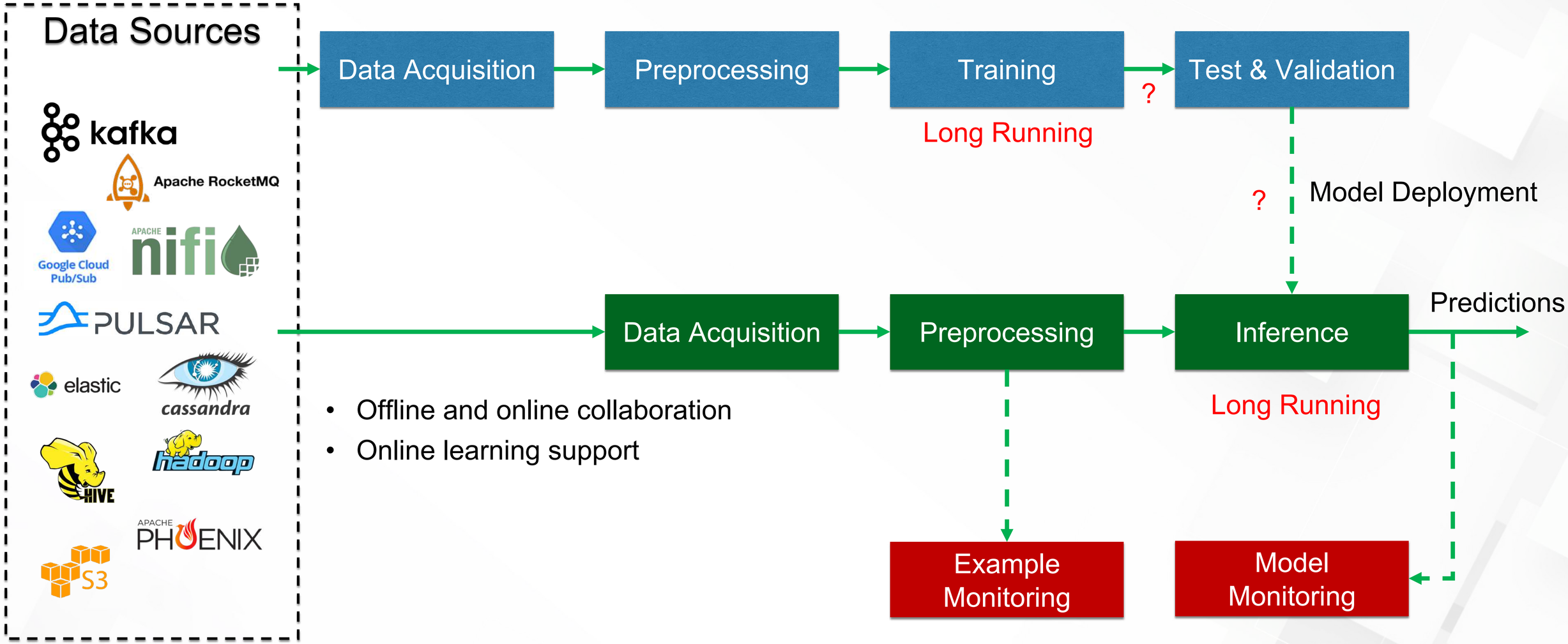
Job status based scheduling

Challenges in AI Workflow

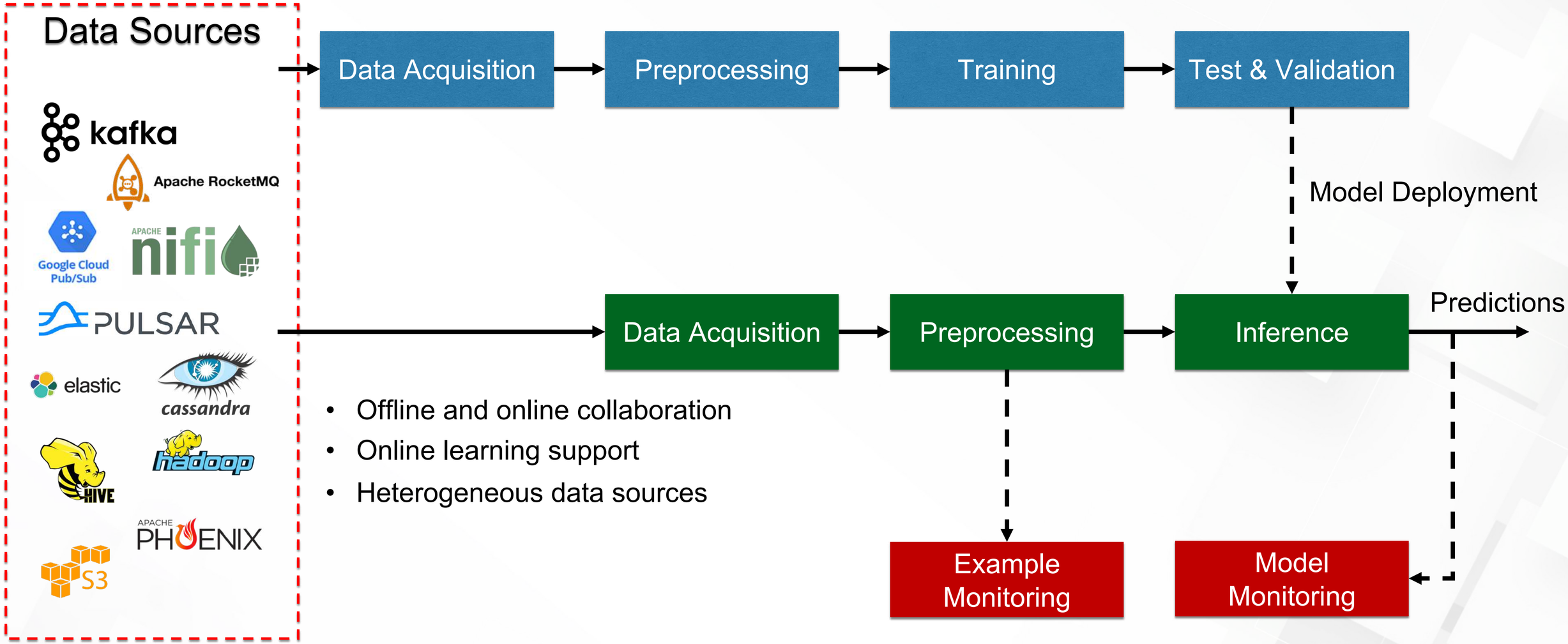


Challenges in AI Workflow

Alibaba Cloud | Worldwide Cloud Services Partner

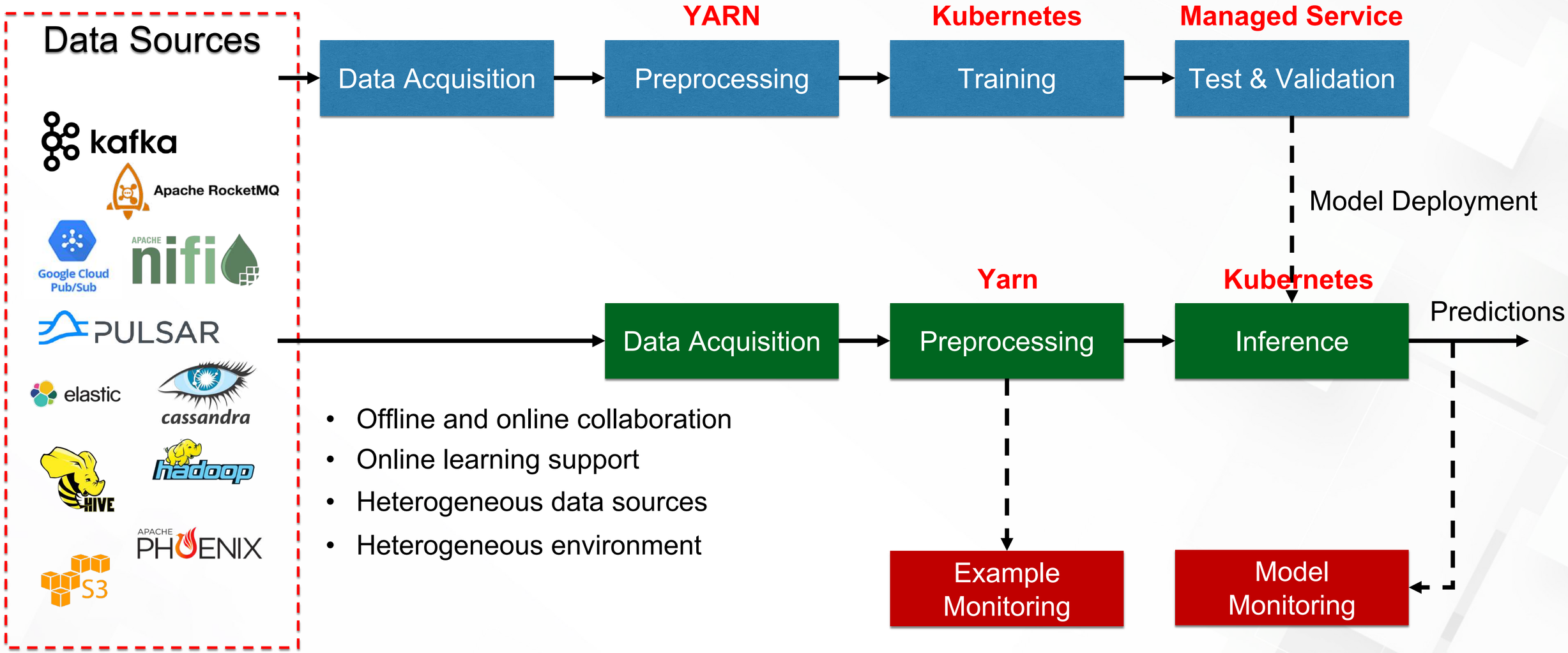


Challenges in AI workflow

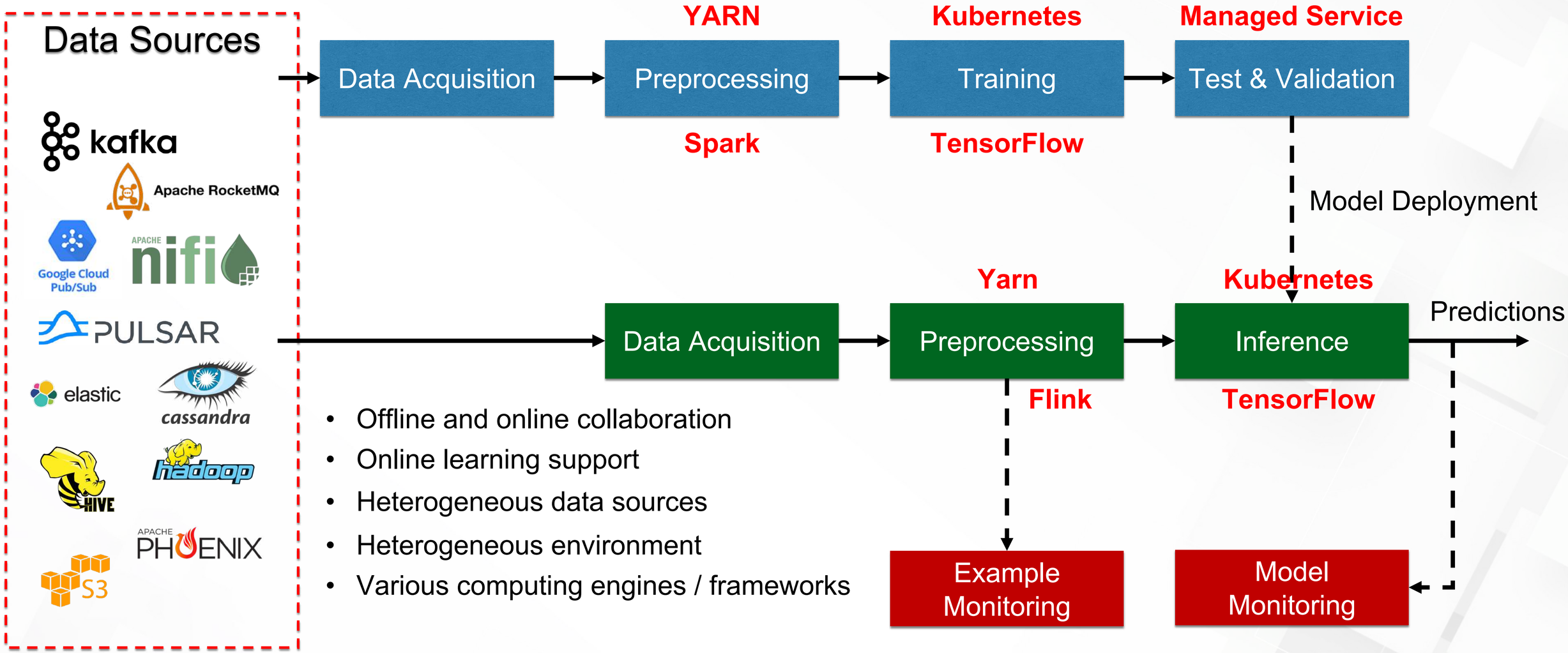


Challenges in AI workflow

Alibaba Cloud | Worldwide Cloud Services Partner



Challenges in AI workflow



Meet AI Flow

- Native support for real-time AI
  - Define workflows with streaming jobs
- Adaptive to various engines
  - Simple Python programs, Flink, Spark, TF, PyTorch, etc.
- Platform agnostic
  - Works with K8S, YARN, Cloud Services, etc

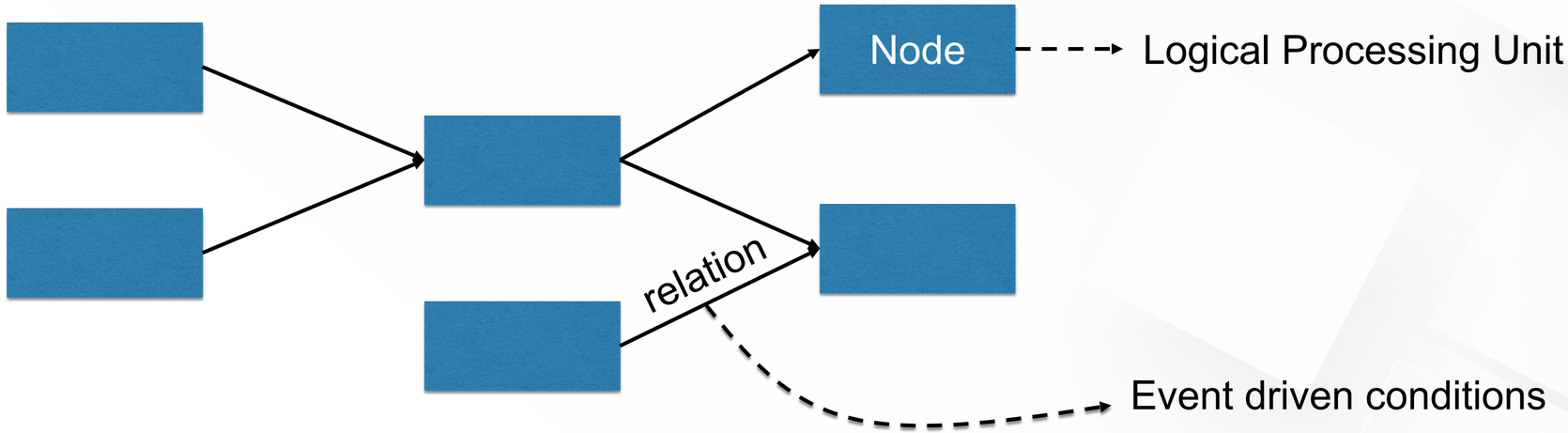
AI Flow Introduction

- Brief Introduction
- Architecture Overview
- Design Walkthrough
- Supporting Services

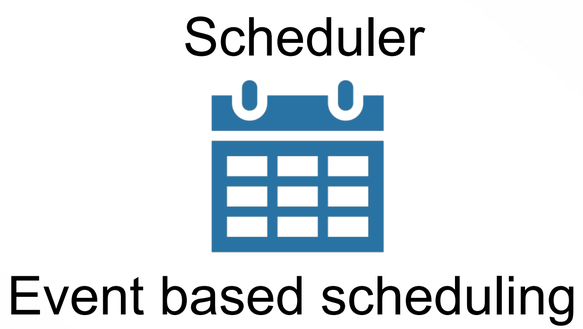
Brief Introduction

- A top level workflow abstraction for big data and AI (esp. real-time AI)

Workflow Definition



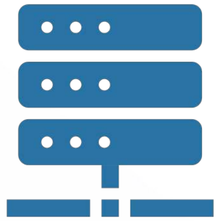
Workflow Execution



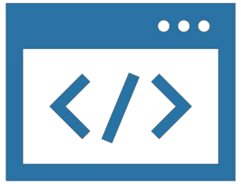
Brief Introduction

- A few supporting services helping fulfill the semantics
  - Metadata Service
  - Notification Service
  - Model Center

Metadata Service

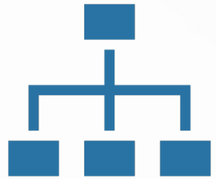


Dataset



Project

Workflow & jobs

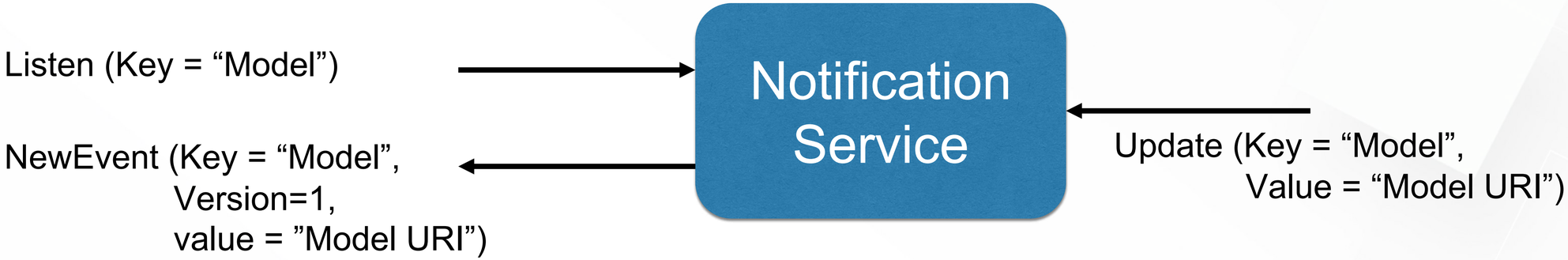


Model Lineage

Artifacts

Notification Service

- Keyed events and event listeners



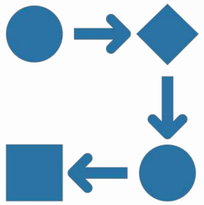
Model Center

Alibaba Cloud | Worldwide Cloud Services Partner

Versioning

Parameters

Metrics



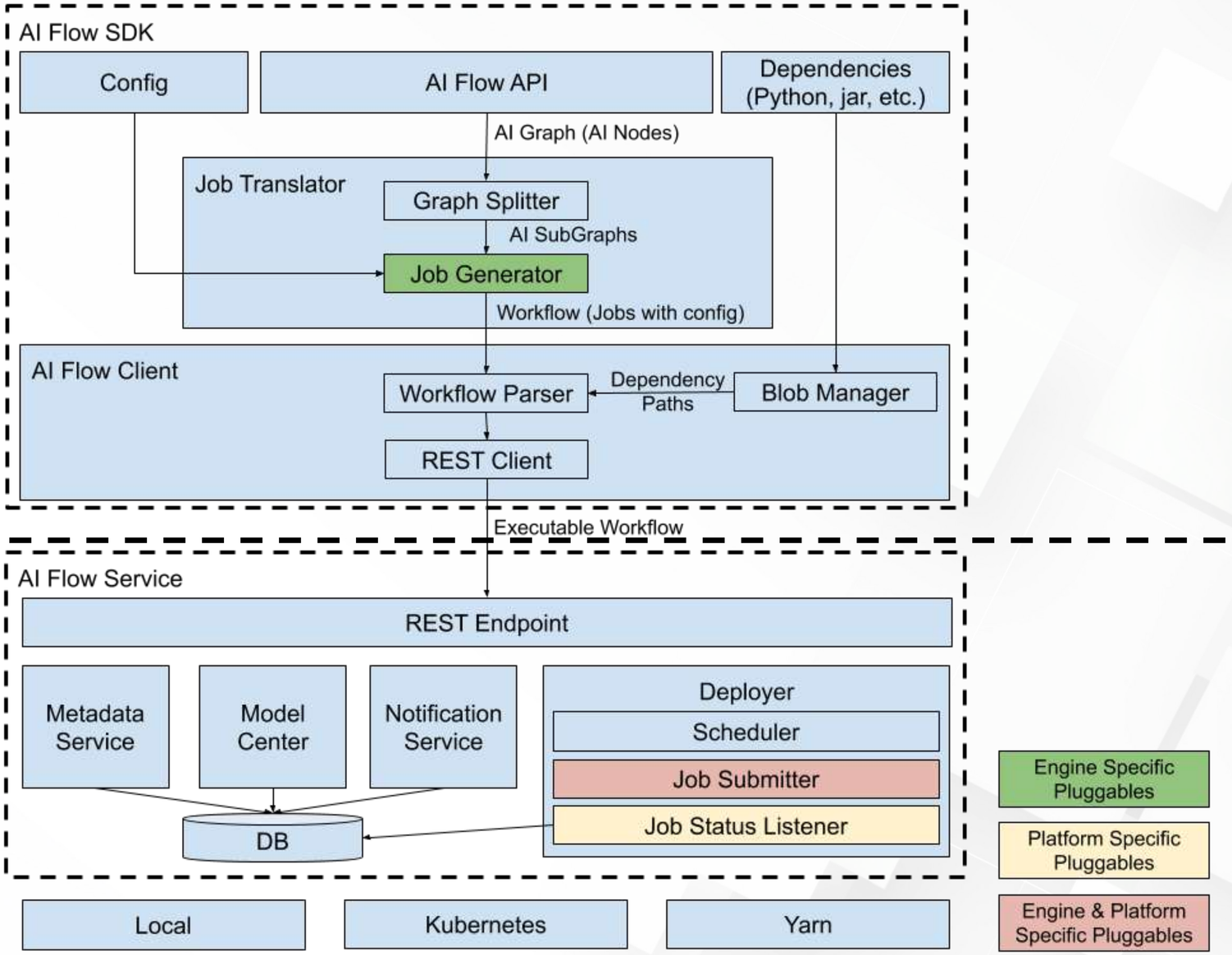
Lifecycle

Visualization

Architecture Overview

Define & Compile

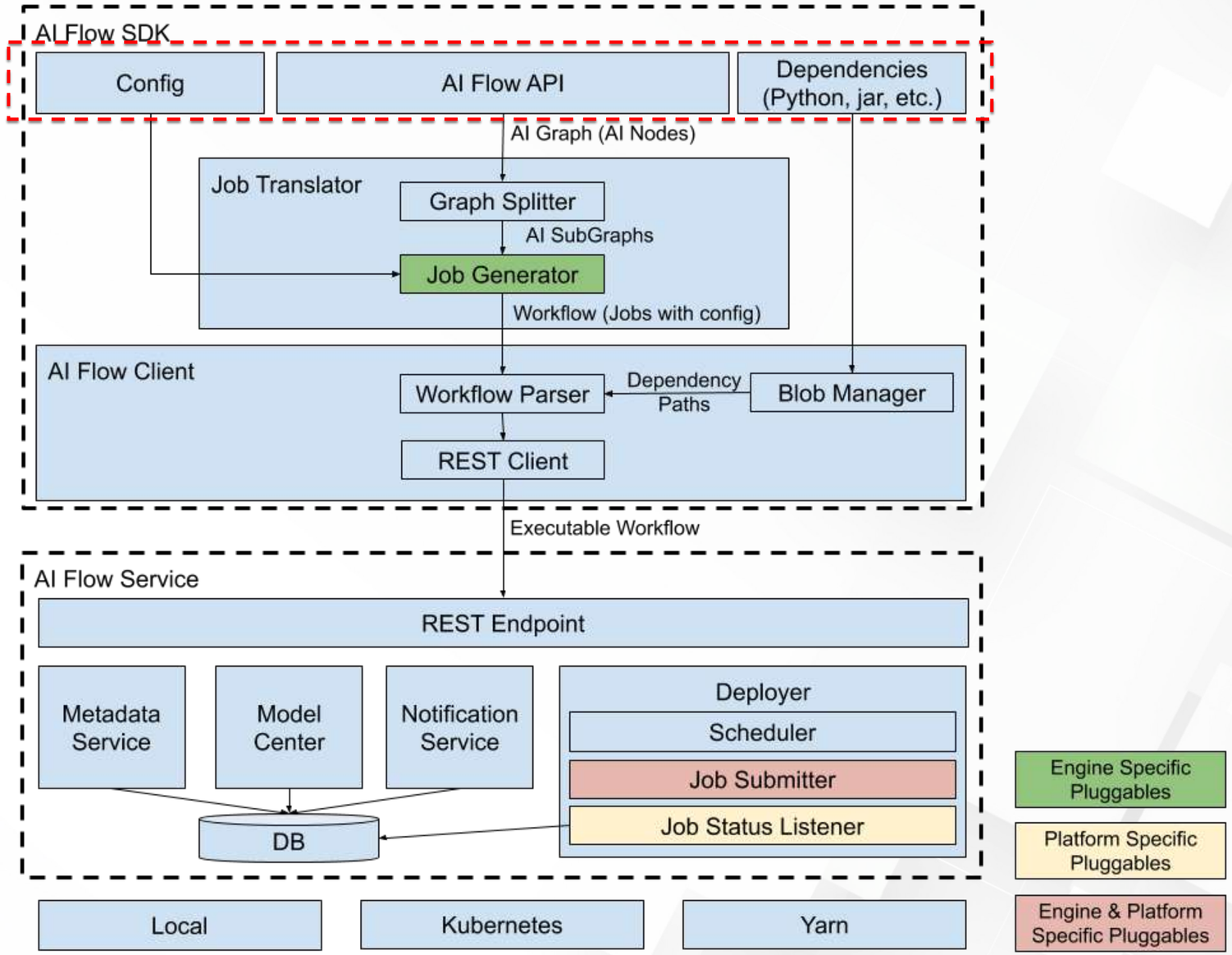
Execute



API Abstraction

Define an **AI Graph**

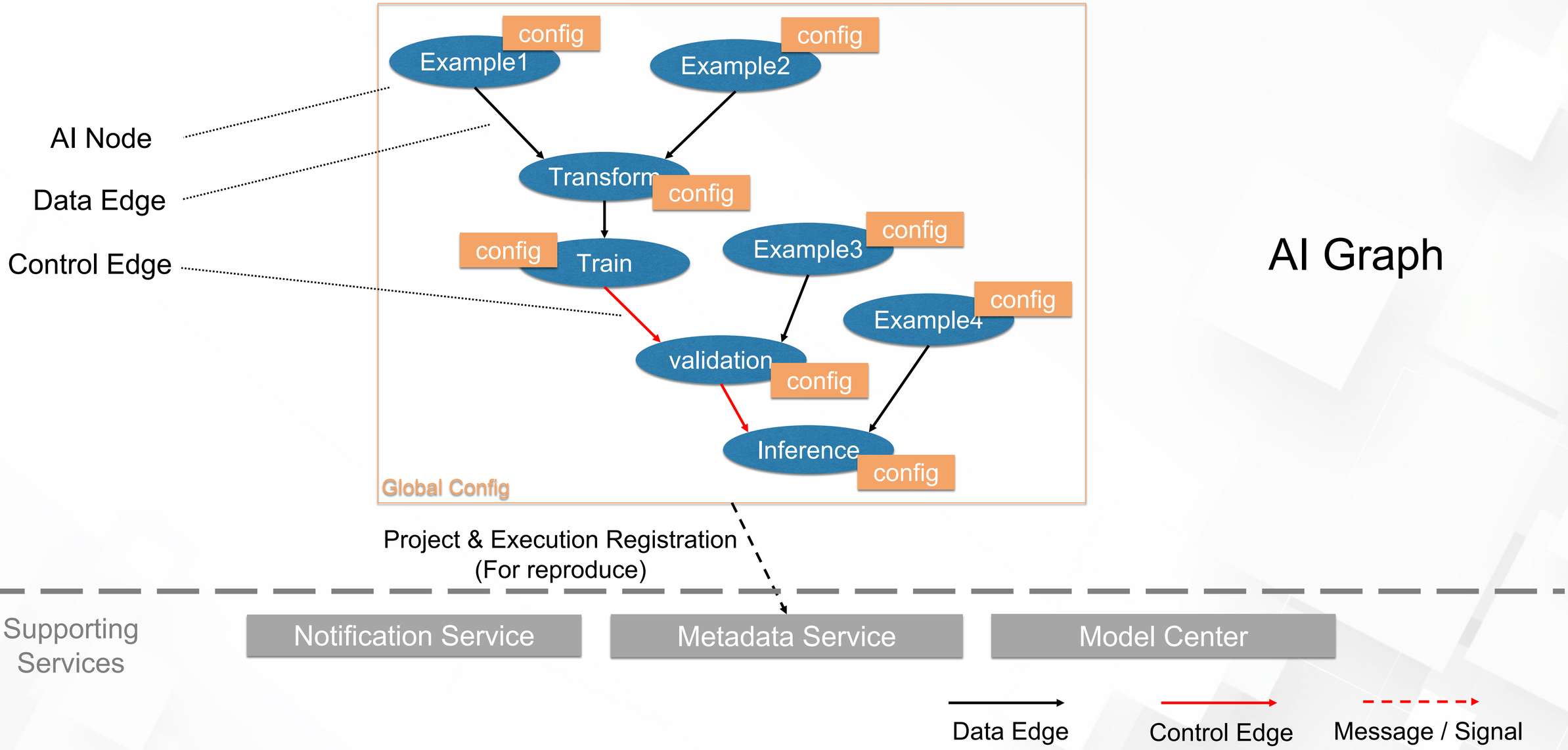
- AI Nodes & relations
- Configs
- Dependencies



Nodes and Relations

- AI Nodes
  - Named logical processing unit with multiple inputs and outputs
  - To be interpreted by the ***Job Generator***
- Relation
  - Data dependency (intra-job)
    - Data flows to be interpreted by the ***Job Generator***
  - Control dependency (inter-job)
    - Event driven control flows to be interpreted by the ***Scheduler***

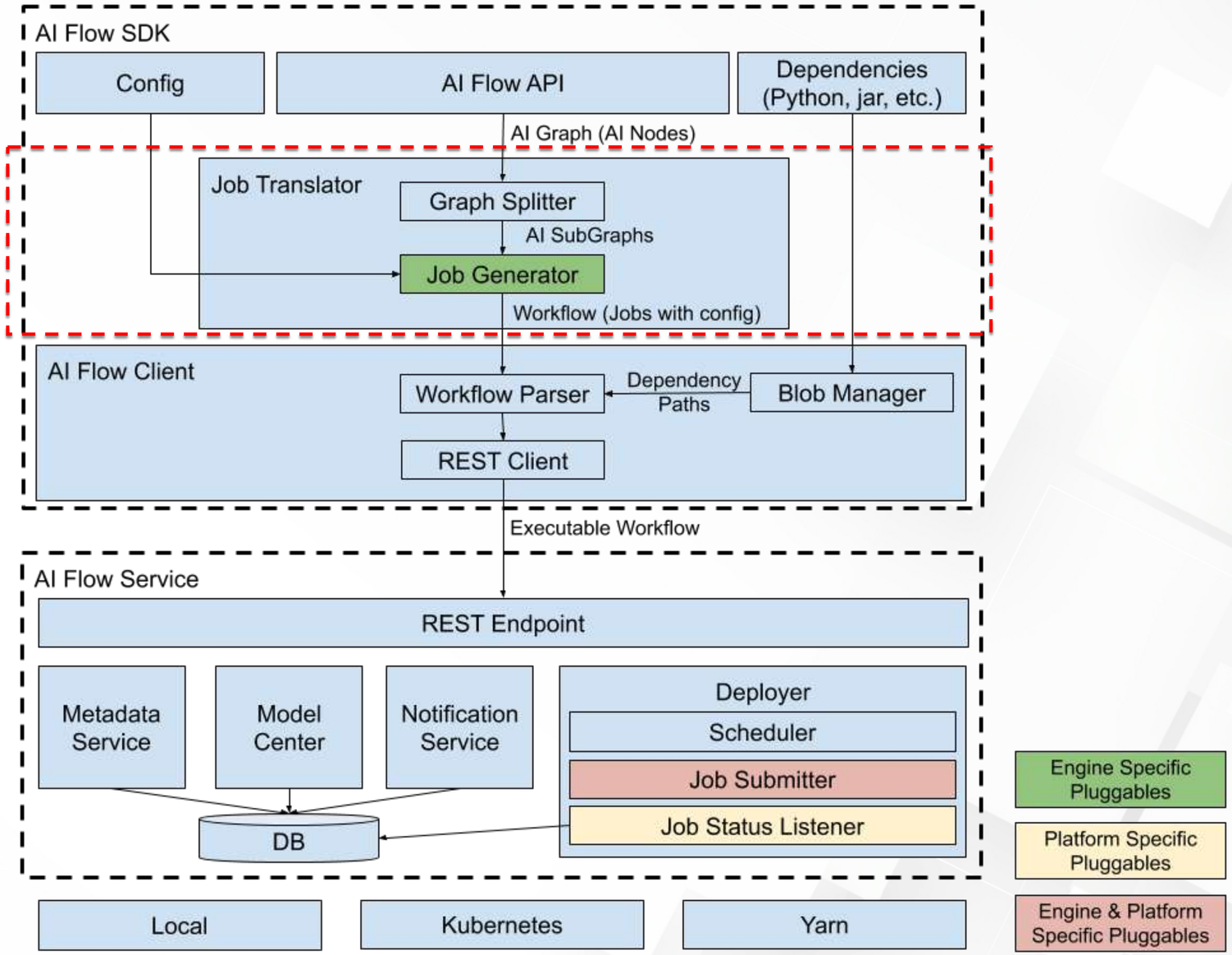
AI Graph Description



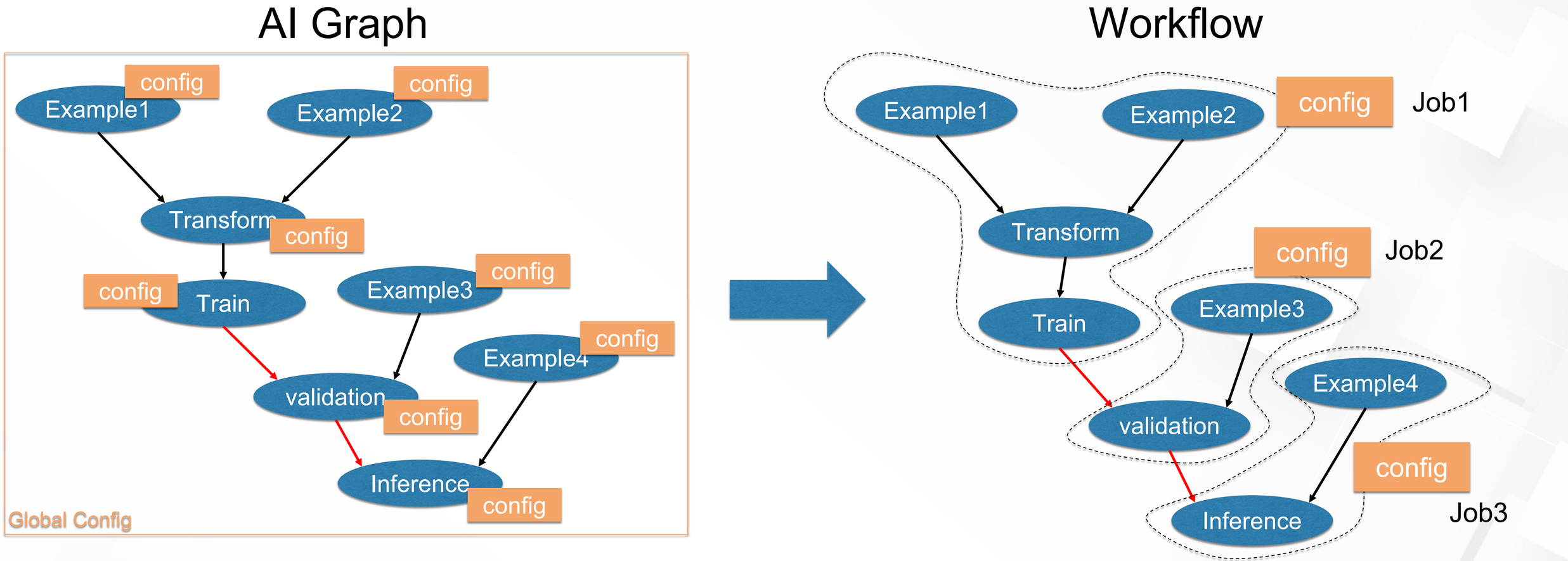
Workflow Compilation

AI Graph → Workflow

- *Split*: AI Graph → Subgraphs
- *Translate*: Subgraph → Job



Workflow Compilation



Execution Blocks

- Specify computing engine and config with execution blocks

```
with af.global_config(af.BaseJobConfig(platform="K8S", engine="Flink", properties={"aa": "aa"})):
    with af.job_config(config=config1):
        ....
    with af.engine('python'):
        with af.job_config(config=config2):
            input = af.read_example(input, …)
            processed = af.transform(input_data_list=[input],
                                     executor=PythonObjectExecutor(python_object=T1()))
            write = af.write_example(input_data=processed, …)
    with af.job_config(config=config1):
        ....
```

Global Config

Flink Job

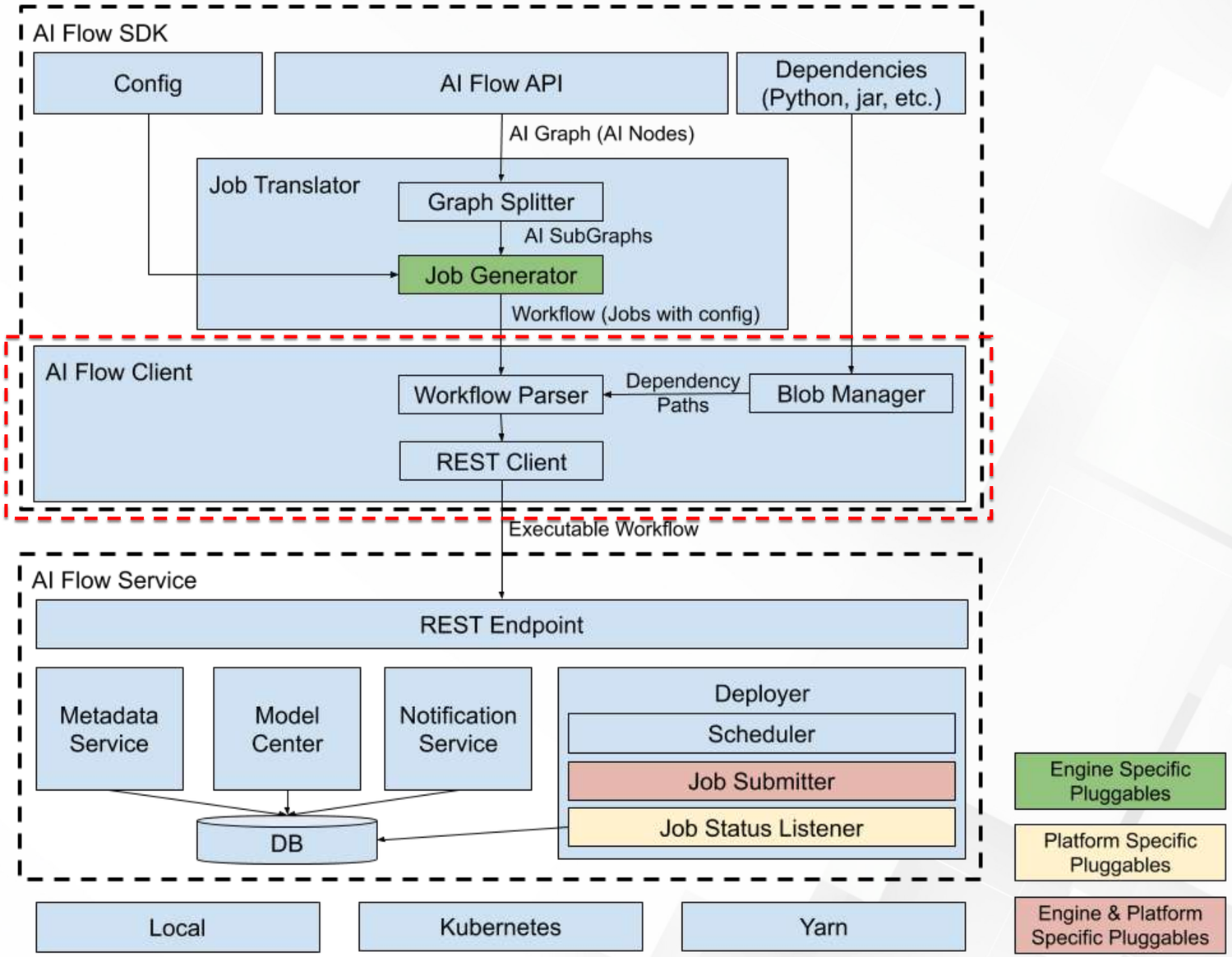
Python Job

Flink Job

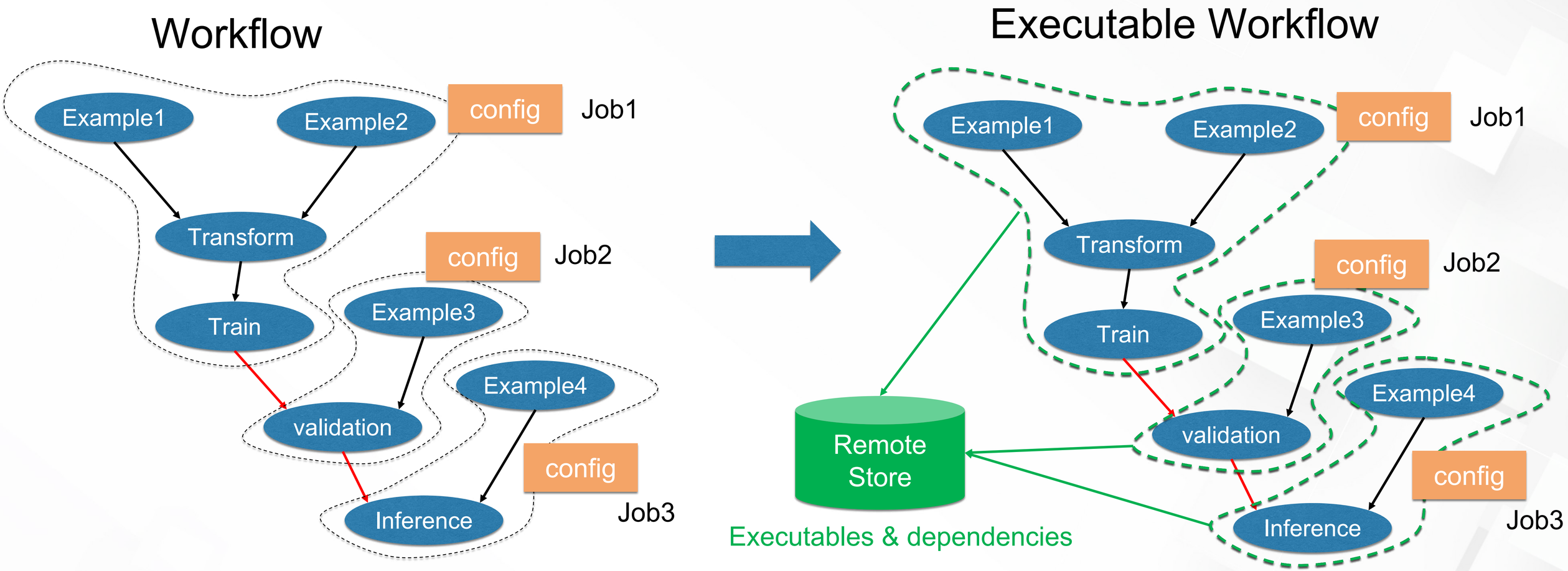
Construct Executable Workflows

Workflow → Executable Workflow.

- Upload code & dependencies
- Update **Workflow** with URI.



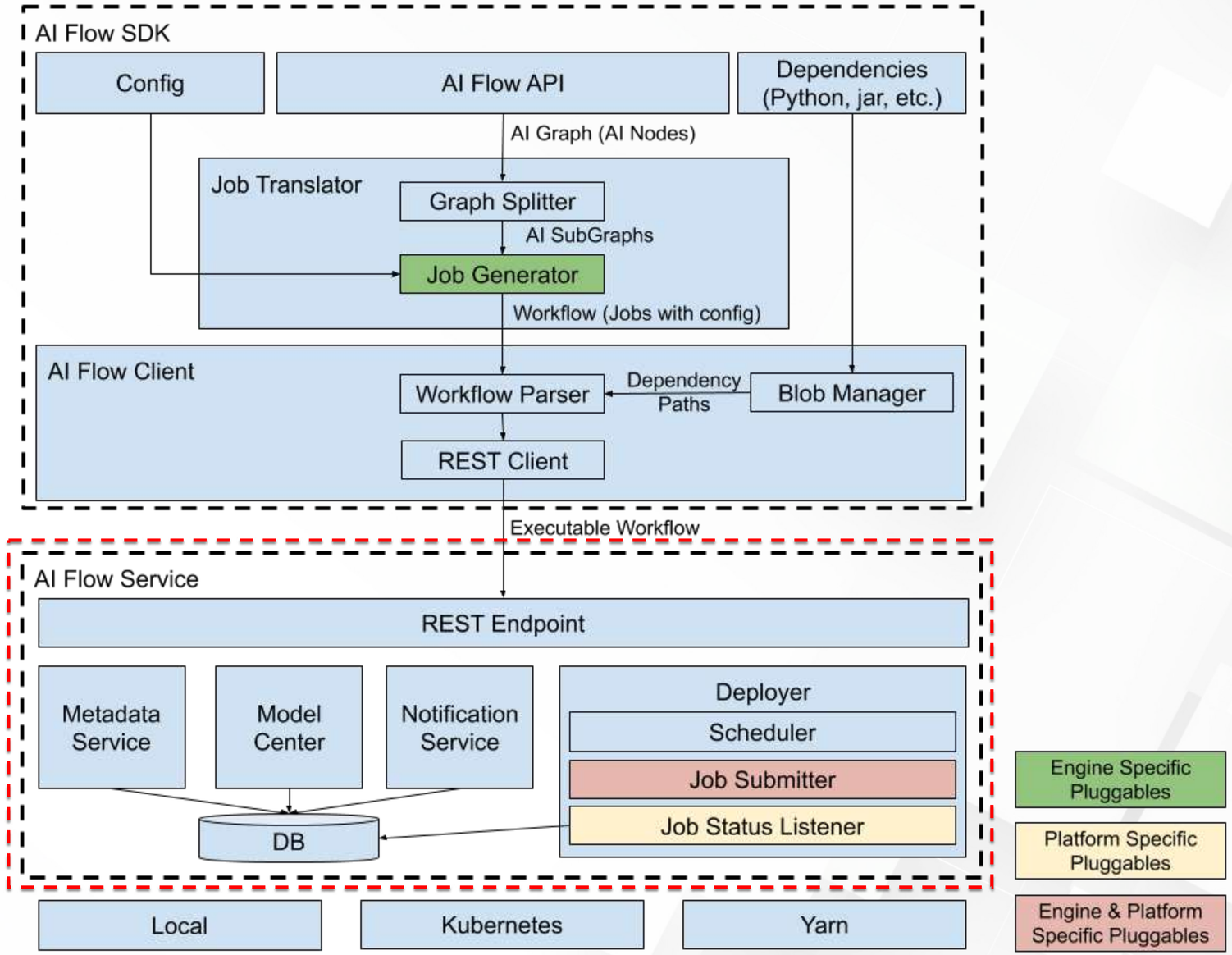
Executable Workflow Construction



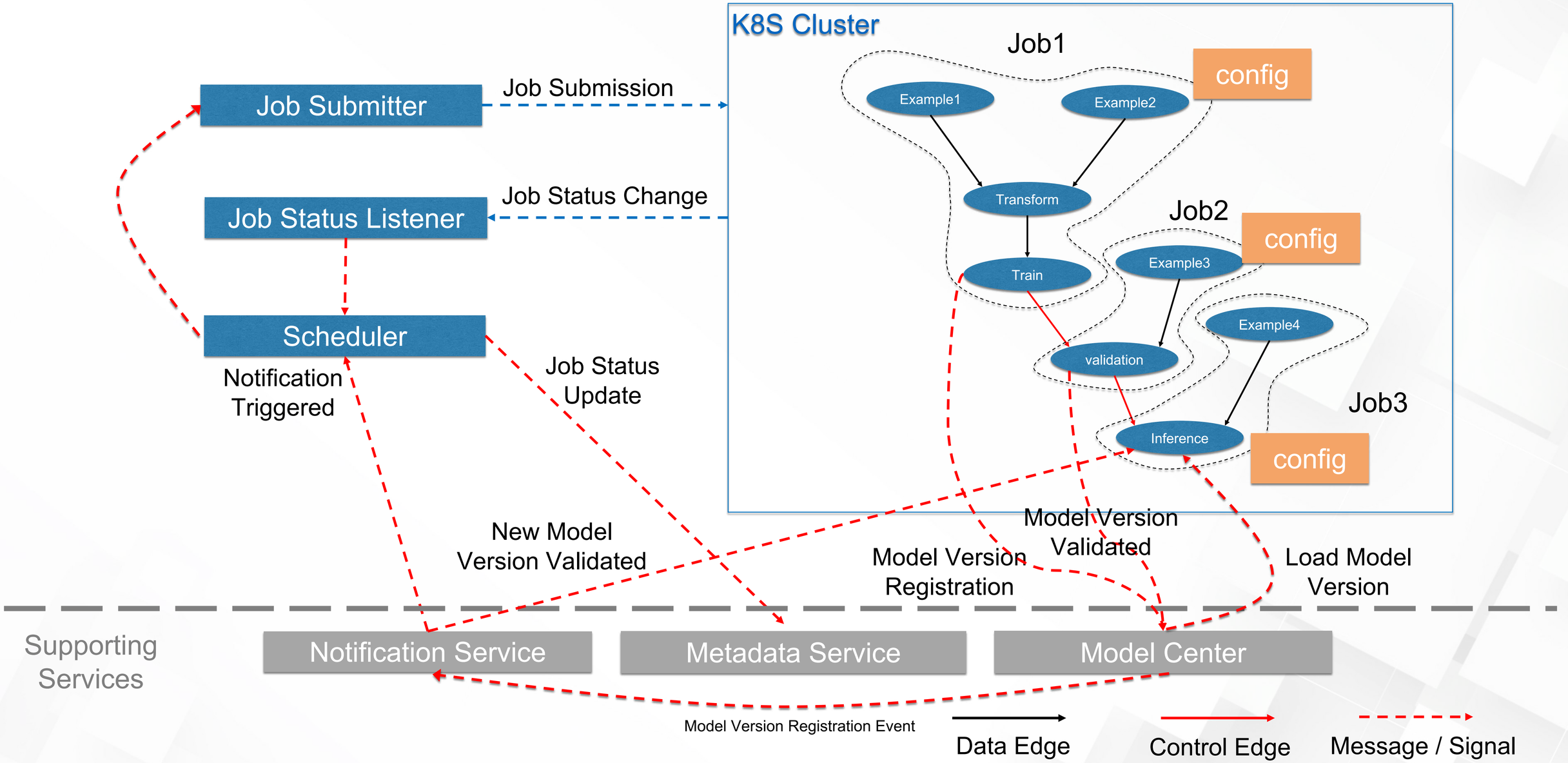
Workflow Execution

Run **Executable Workflows**

- Schedule jobs
- Submit jobs to different platforms



An Example



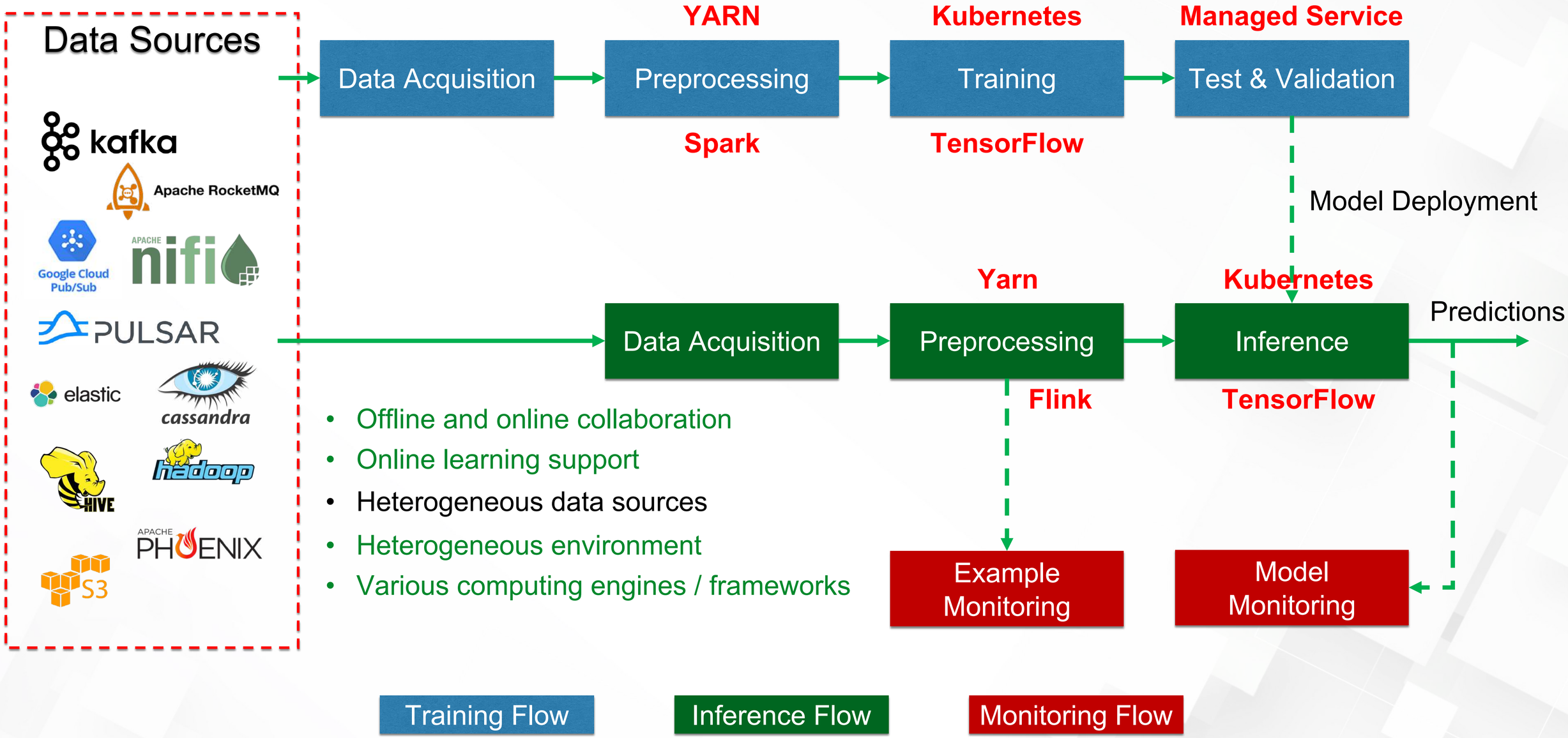
Flink AI Flow

- Flink AI Flow
- Flink AI Flow & Flink ML Pipeline
- Flink AI Flow & PyFlink

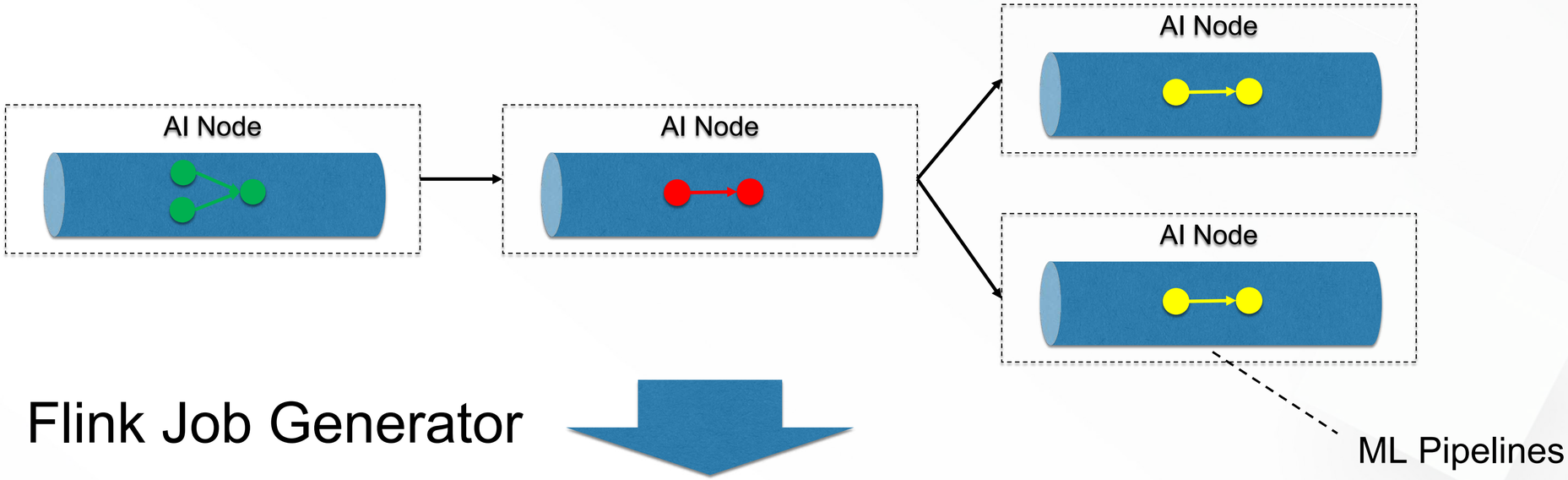
Flink AI Flow

- Implementation of AI Flow semantics
  - Flink Job Generator
  - Flink Job Submitter (Local / K8S)
  - ~1000 lines of code (easy to integrate)
- Work seamlessly with Flink Ecosystem
  - PyFlink
  - Flink ML Pipeline
  - TF on Flink

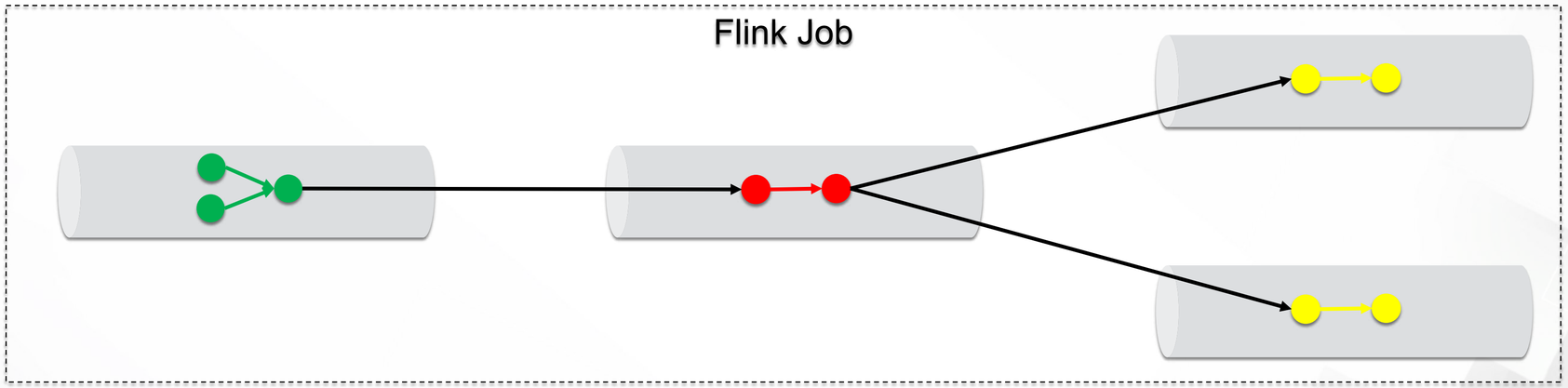
Flink AI Flow



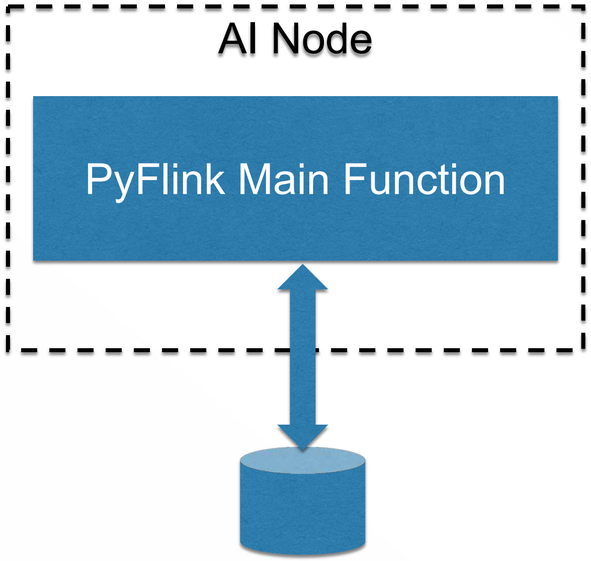
Flink AI Flow & ML Pipeline



Flink Job Generator



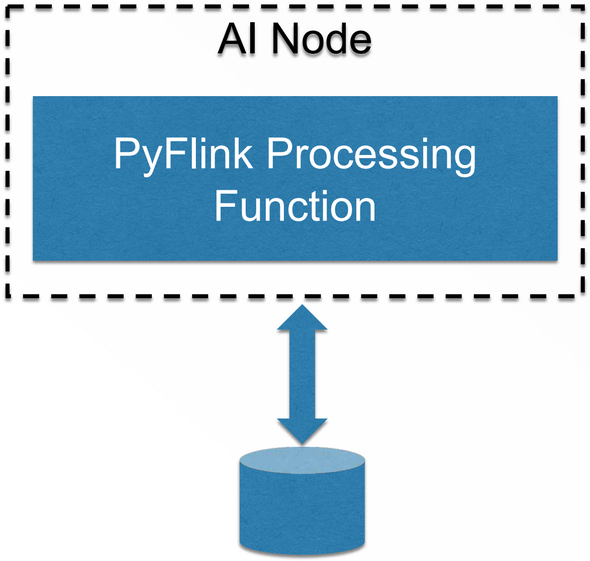
Flink AI Flow & PyFlink



engine = Python

As an ordinary Python function

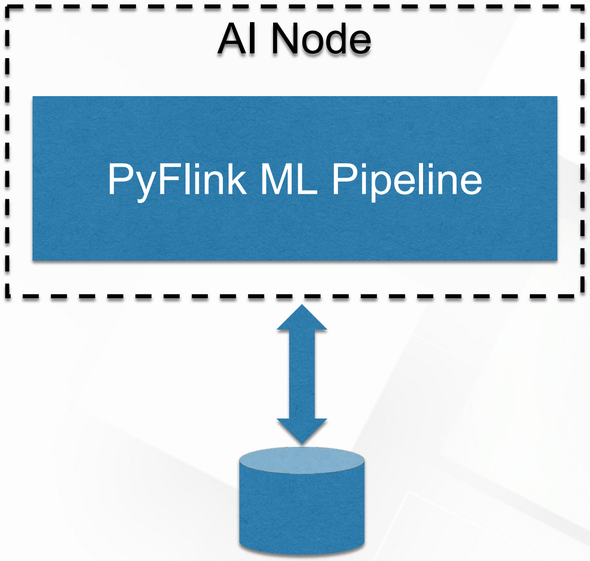
(Users deal with IO)



engine = Flink

As chainable PyFlink Processing Functions

(Flink AI Flow deals with IO)



engine = Flink

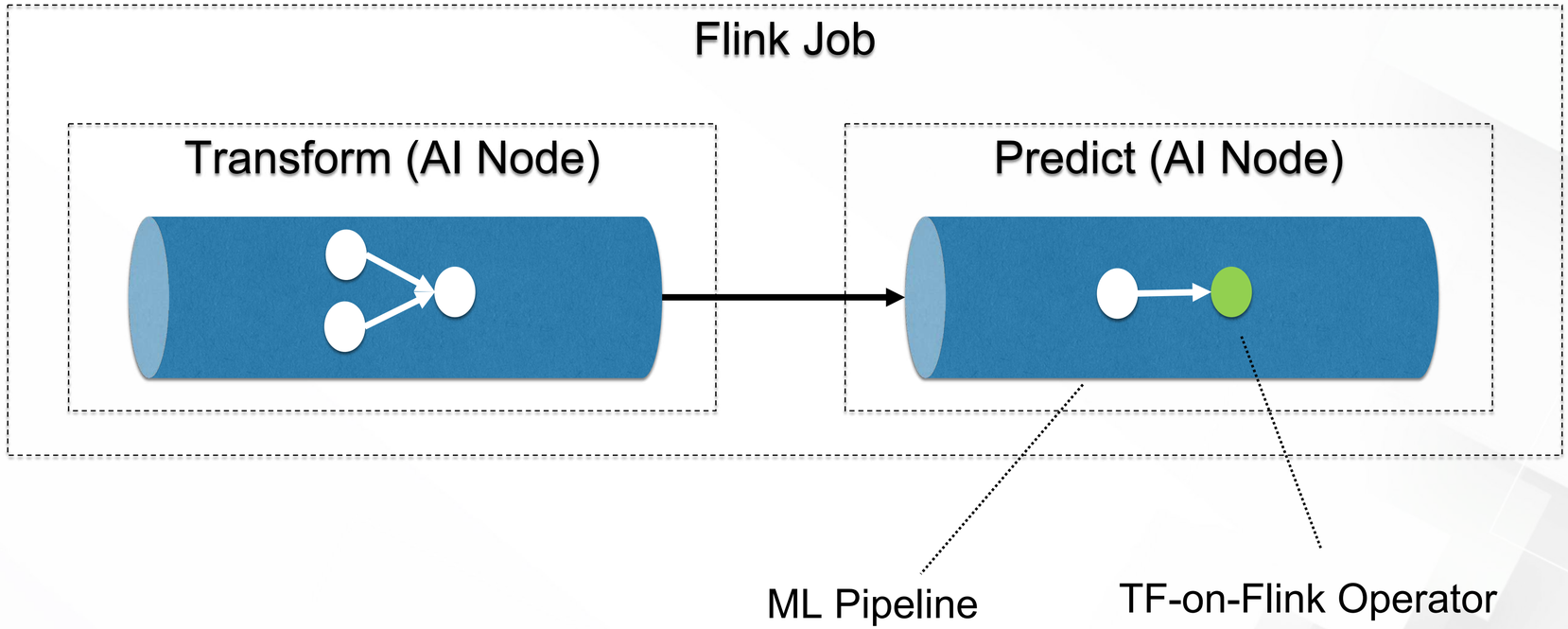
As PyFlink ML Pipeline Stages

(Flink AI Flow deals with IO)

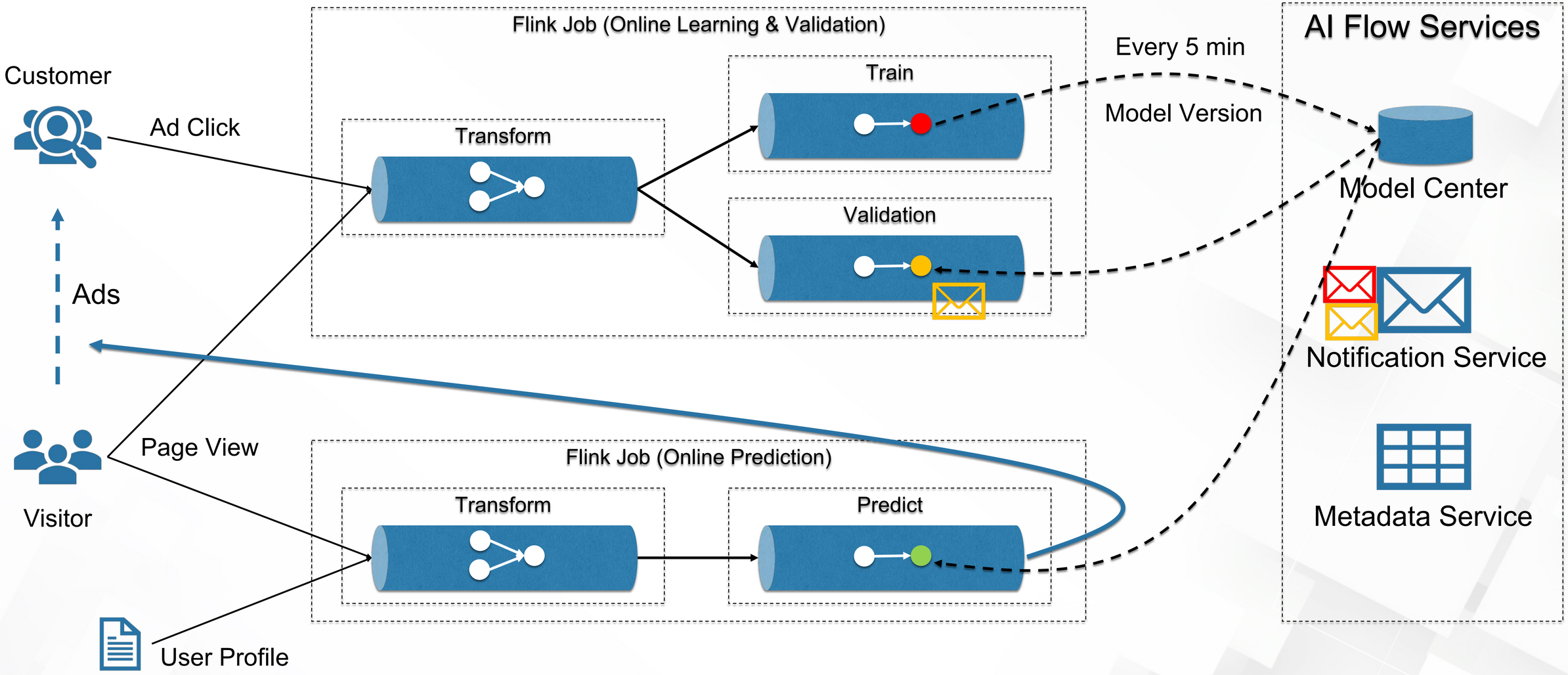
Available after Flink 1.11

Flink AI Flow & TF-on-Flink

- Recap for TF-on-Flink
  - TensorFlow as a Flink Operator
  - No intermediate storage between Flink and TF



An Example



Summary

Summary – AI Flow

- A workflow with streaming job support
- Supporting services to facilitate AI use cases
- Works perfectly well with Apache Flink
  - Also adaptive to other engines

Alibaba Cloud | Worldwide Cloud Services Partner

AI Flow + Apache Flink
AI made easy!

Open Source Soon!

WE ARE HIRING!!!

Summer Interns: Scan QR Code

Full-Time: Send resume to jiangjie.qj@alibaba-inc.com

Alibaba Cloud |

Worldwide Cloud Services Partner