
Flyte

Platforma za orkestracijo

procesov

(github.com/flyteorg/flyte)

Klemen Špruk

Kaj je Flyte

- Flyte je odprtokodna platforma za orkestracijo delovnih procesov, zasnovana za zanesljivo in razširljivo obdelavo podatkov in strojno učenje.



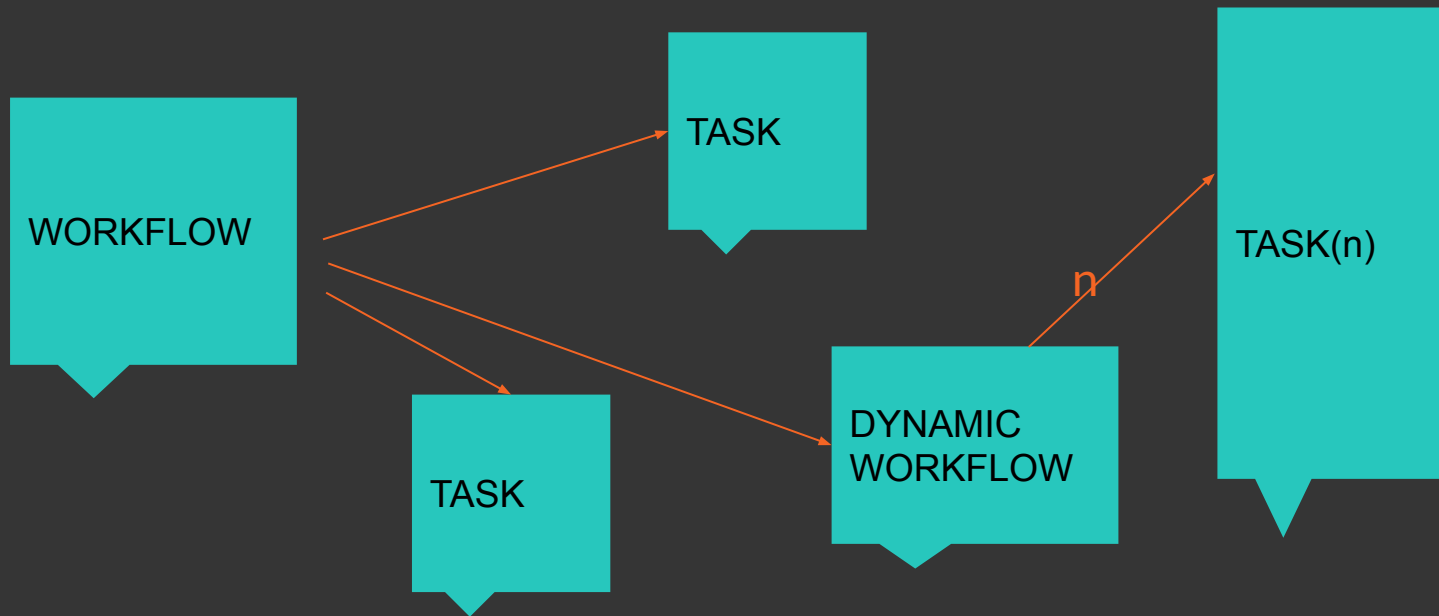
FFAI & DATA
GRADUATE PROJECT



Prednosti

- **Modularna arhitektura omogoča ponovno uporabo kode.**
- **Izboljšanje produktivnosti z avtomatizacijo procesov**
- **Natančna sledljivost vseh izvedb.**
- **Podpora za kompleksne delovne procese z zankami, pogojnimi logikami in dinamičnimi delovnimi tokovi.**

Kako Flyte deluje?



Workflow (DAG - directed acyclic graph)

Directed - Usmerjen: smer potovanja po grafu je dolocena

Acyclic - Aciklicen: ni ponavljajocih zank ali ponavljanj, s tem je zagotovljeni da izvajanje nalog poteka naprej in se zakljuci

Flyte Workflow je kombinacija nalog urejenih v usmerjen aciklicen graf. Z definicijo workflowa deklarativno definiramo DAG ki ga Flyte izvede.

Prednosti:

- Vzporedno izvajanje neodvisnih nalog
 - Izolacija napak (ce se v enem robu grafa zgodi napaka, se drug rob vseeno izvede - definiramo "failure policy")
 - Modularnost - ponovna uporaba nalog v "podrejenih" workflowih
-



Arhitektura v Flyte

- Podatkovna baza
- Podatkovno skladišče
- Kubernetes cluster
 - ◆ Ima definirane meje za stevilo CPU, kolicino RAMa
 - ◆ Kubernetes samodejno povečuje/zmanjšuje stevilo enot za procesiranje

Primeri uporabe:

- Obdelava podatkov in ETL (Extract, Transform, Load)
- Strojno učenje in umetna inteligenca
- Obdelava prostorskih podatkov
- Izvajanje finančnih analiz
-

Integracije Flyte

Programski jeziki: Python, Java, Scala, JavaScript

Ponudniki “cloud” infrastructure: AWS, Google Cloud, Azure

Podatkovna skladišča: AWS S3, Google Cloud Storage, Azure Blob Storage

Strojno učenje in AI: TensorFlow, PyTorch,

—

Moji primeri uporabe:

- procesiranje lidar podatkov,
- procesiranje podatkov satelitskih zajemov,
- izdelava definicije osnovnega OS preko Docker platforme, kjer se potem poganja algoritem obdelave podatkov
- “scalable deployment” algoritma za obdelavo podatkov
- analiza uspešnosti obdelave podatkov (ELK) .

Primer

HVALA ZA POZORNOST!

Q&A