

# 開源軟體介紹 - 以MLOps為例

葉信和 / Hsin-Ho Yeh  
Software Engineer / CEO @ 信誠金融科技  
[hsinho.yeh@footprint-ai.com](mailto:hsinho.yeh@footprint-ai.com)



2023/03/22

# Download Slides

<http://bit.ly/3JA2scn>



# About me

- 2020 - Present at 信誠金融科技
  - **Shrimping**: A data-sharing platform
    - <https://get-shrimping.footprint-ai.com>
  - **Tintin**: a machine learning platform for everyone
    - <https://get-tintin.footprint-ai.com>
  - **KaFeiDo**: machine learning platform for green economy
- 2016 - 2020 at IglooInsure (16M+ in series A+ 2020)
  - Provide digital insurance for e-economic world
  - Funded in KUL, Headquartered in Singapore
  - First employee/ Engineering Lead / Regional Head/ Chief Engineer
- 2013 - 2016 at Studio Engineering @ hTC
  - Principal Engineer on Cloud Infrastructure Team
- 2009 - 2012 at IIS @ Academia Sinica
  - Computer vision, pattern recognition, and data mining
- CS@CCU, CS@NCKU alumni



# Agenda

- What is Open Source?
- How to contribute to Open Source?
- Kubeflow Case Study
- FAQ

## What is Open Source Software?

“Open-source software is a type of computer software ... the copyright holder grants users the rights to **use**, **study**, **change**, and **distribute** the software to anyone and for any purpose.” [1]

[1] [https://en.wikipedia.org/wiki/Open-source\\_software](https://en.wikipedia.org/wiki/Open-source_software)

# What is Open Source Software? (1/4)

	<b>Open Source Software</b>	<b>Proprietary Software</b>
Ownership	Everyone	A group of limited people
Usability	Difficult use for unskilled users Ex: Desktop Linux	Feature neat to end-user Ex: Microsoft Windows
Security	Better (?) Ex: <u>PHP source code is injected backdoors by hackers</u>	Less
Stability	Less	Better
Opacity	Transparent	Closed

### The misunderstanding on open source software

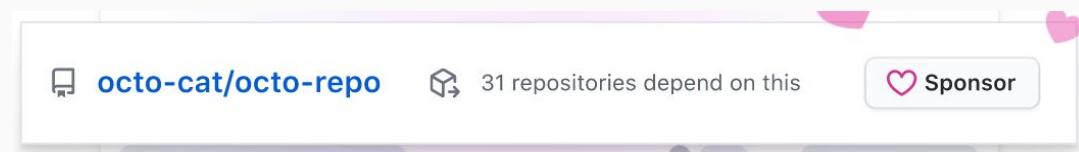
- Open source software is **Free** in terms of Freedom, not Monetary.
- Open source developer is **NOT** a free labor [1]
  - Sponsorship by companies
  - Add values( a.k.a layers) on the existing project
  - Professional consult
- Open source means **Pool** quality
  - Good quality when the project gained enough eyeballs
  - On average the density of code defect is smaller than proprietary software [2]

[1] <https://softwareengineering.stackexchange.com/questions/100685/making-money-with-open-source-as-a-developer>

[2] <https://blogs.worldbank.org/opendata/quality-open-source-software-how-many-eyes-are-enough>

### Where open source communities lives?

- Github
  - Free for any public repositories
  - Github Sponsor



# Key files to understand a github project

- README.md
  - How to build and use the project
- LICENSE
  - Terms and Conditions
- CONTRIBUTING.md
  - Rules on contributing
- ROADMAP.md
  - Project roadmap
- CODE OF CONDUCT.md
  - Guidance on rule, responsibilities

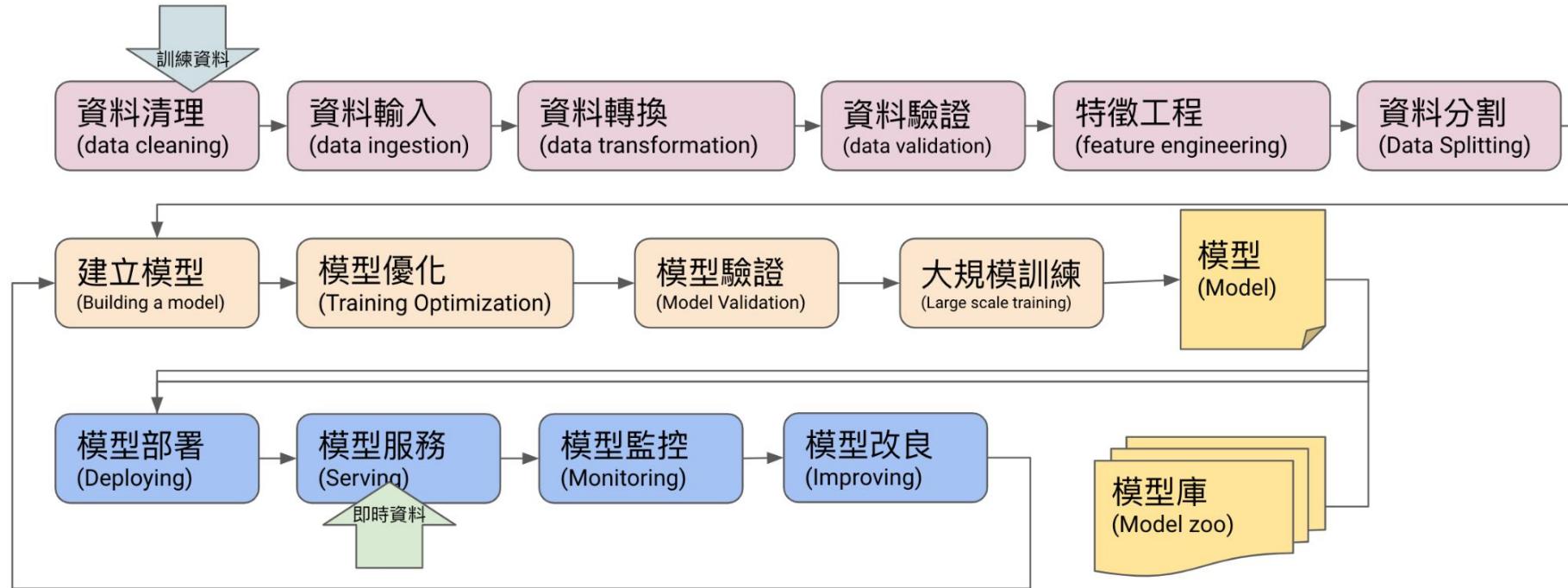
What is kubeflow?



Kubernetes + ML  
= Kubeflow

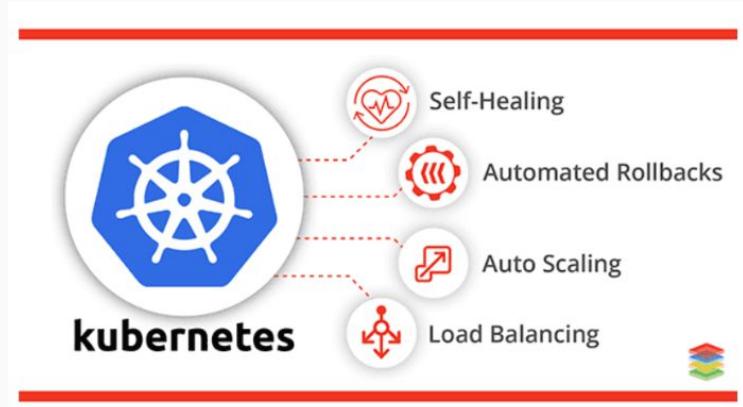
The Kubeflow project is dedicated to making deployments of machine learning (ML) workflows on Kubernetes simple, portable and scalable.

# Real-world Machine Learning Application - End-to-End ML LifeCycle



# Why machine learning on Kubernetes?

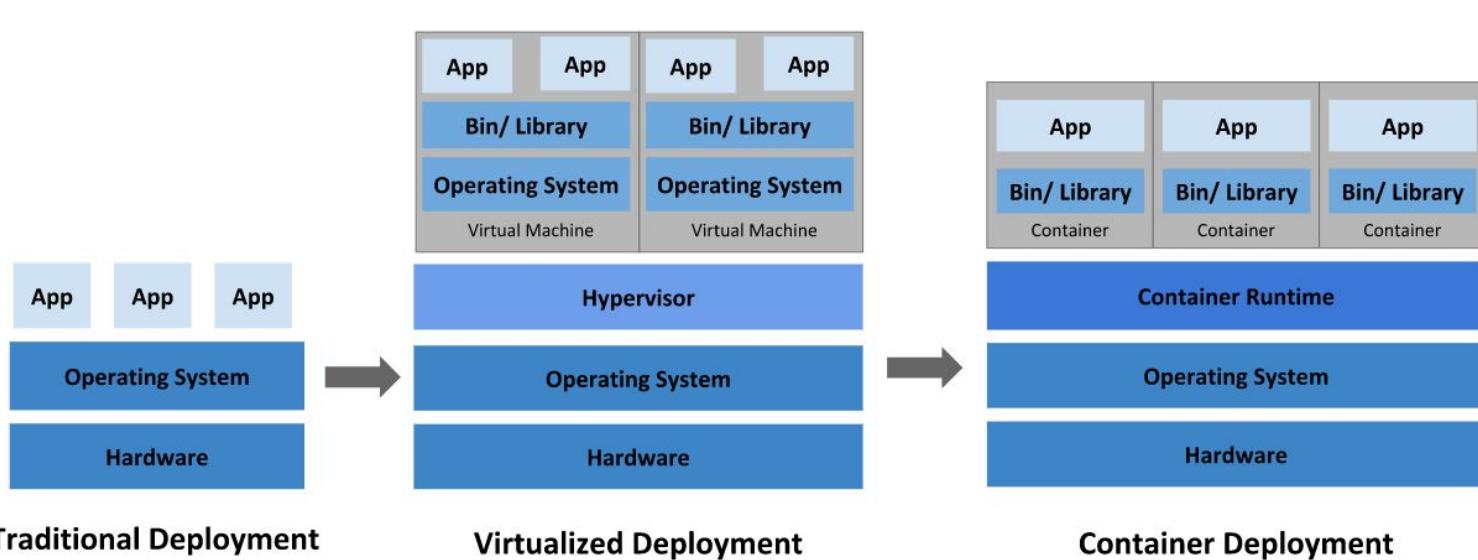
- Portability
  - Dev/Staging/Prod
  - Laptop/Edge/Cloud environment
- Scalability
  - Hyperparameter tuning, production workloads
- Isolation
  - Workloads from different experiments



# What is container?

- Container

- Container Image = Application code + dependencies
- Runtime environment (cgroups, namespaces, env vars)



# History Of Kubernetes

- Borg: the predecessor to Kubernetes
  - Google revealed the first time of its detail in an academic research paper, describing a “cluster manager that runs hundreds of thousands of jobs, from many thousands of different applications, across a number of clusters each with up to tens of thousands of machines.”[1]
  - A in-house cluster manager system inside Google for running every google services including Gmail, Google Maps, Google Docs...[2]
  - In a scale with ‘over 2 billion containers per week’ [3]
- The very first version of Kubernetes was released in 2015
- The latest version is v1.26, released at 2023.



[1] <https://research.google/pubs/pub43438/>

[2] <https://www.wired.com/2016/04/want-build-empire-like-googles-os/>

[3] <https://cloud.redhat.com/blog/building-kubernetes-bringing-google-scale-container-orchestration-to-the-enterprise>

# Case Study: Open AI

OpenAI adopted Kubernetes since 2016 for portability, cost saving, and improved efficiency[1,2].

Years	Nodes	Estimated Cost [3]
2018	2,500	= 3 * 2500 * 24 = US\$ 180,000 / day
2021	7,500	= 3 * 7500 * 24 = US\$ 540,000 / day
2023	?	

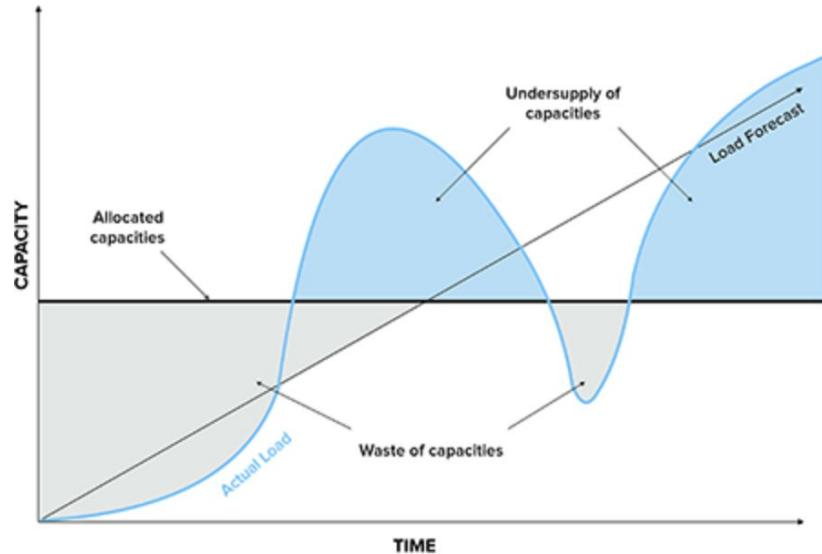
[1] <https://kubernetes.io/case-studies/openai/>  
<https://blog.pichuang.com.tw/20230214-openai-scaling-kubernetes-to-7500-nodes.html>

[2] [https://blog.pichuang.com.tw/20230214-openai-scaling-kubernetes-to-7500-nodes.html#\\_1](https://blog.pichuang.com.tw/20230214-openai-scaling-kubernetes-to-7500-nodes.html#_1)

[3] AWS P3.2xlarge equips 8 vCores, 61 GB Memory, 1 Core of V100-16Gb GPU, charges US\$ 3 per hour.

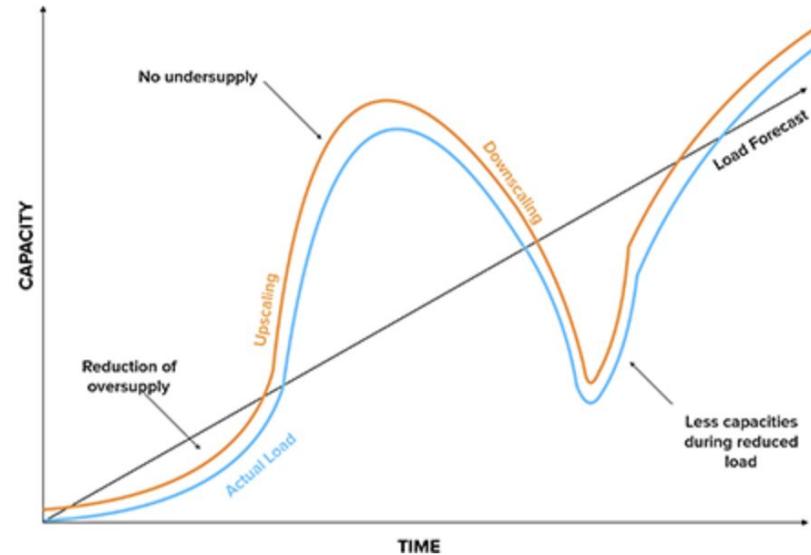
# Cost, Cost, and Cost

Static Architecture



Static architectures are based on estimated load expectancy and are not flexible enough to adapt to unexpected load peaks or lulls.

Auto Scaling Architecture



Auto Scaling is the most cost-efficient solution for a fluctuating load. High performance and thereby user satisfaction are retained at all times.

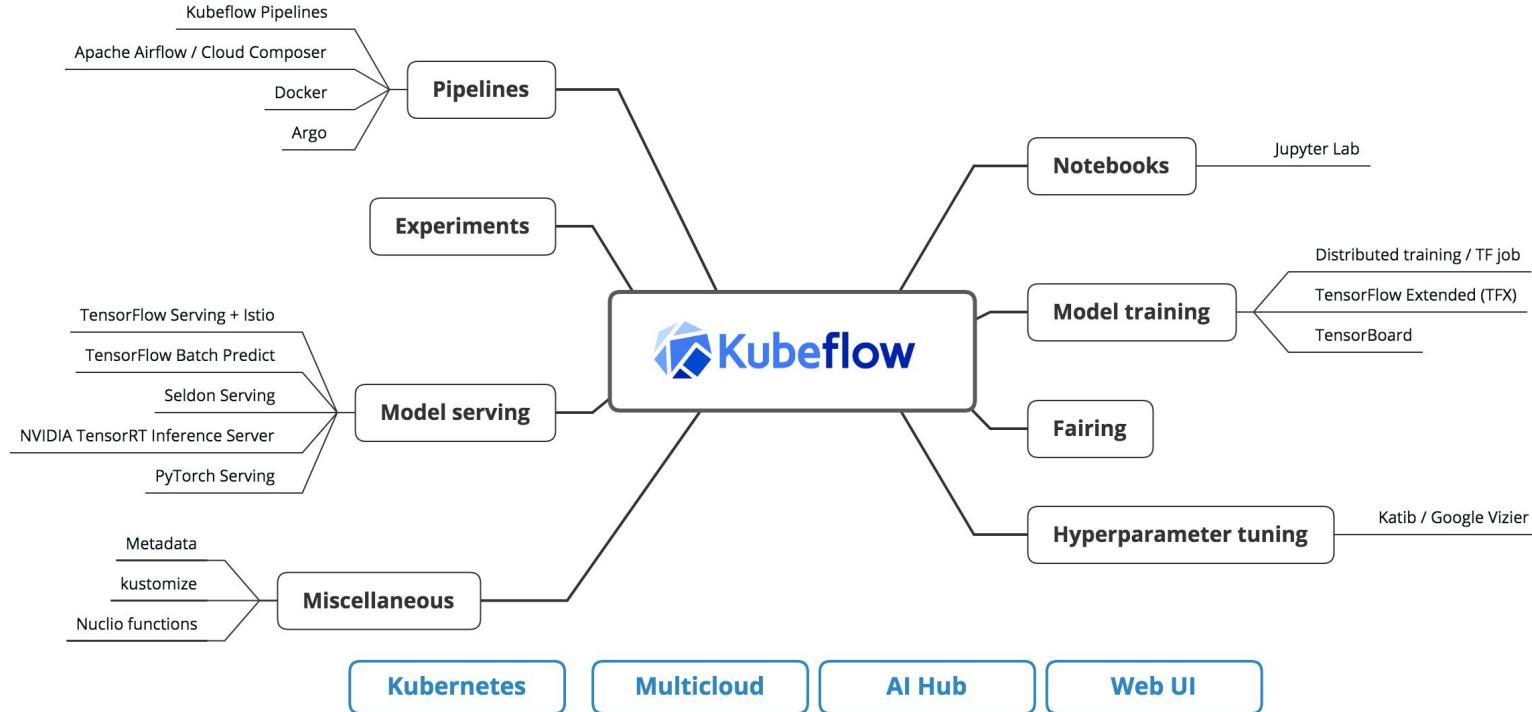
**Oh, you want to use ML on K8s?**

**Before that, can you become an expert in:**

- Containers
- Packaging
- Kubernetes service endpoints
- Persistent volumes
- Scaling
- Immutable deployments
- GPUs, Drivers & the GPL
- Cloud APIs
- DevOps
- ...



# Architectures



# How to contribute Open Source?

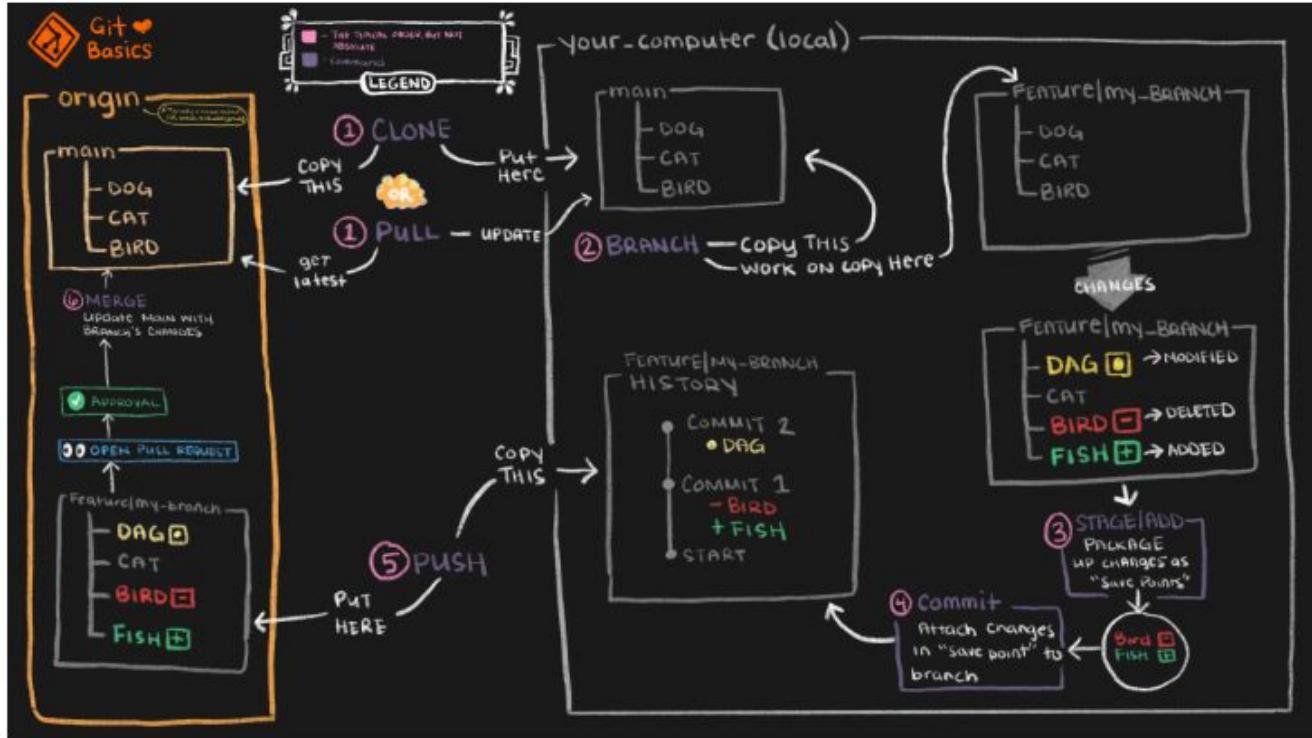
## The two most popular way to contribute

1. Raising an issue where you can feel the project could be improved
  - a. Report a problem/bug (kind/bug)
  - b. Suggest new features (kind/feature)
  - c. Gap in documentation (doc)
2. Contribute to code
  - a. Pull-Request in Github

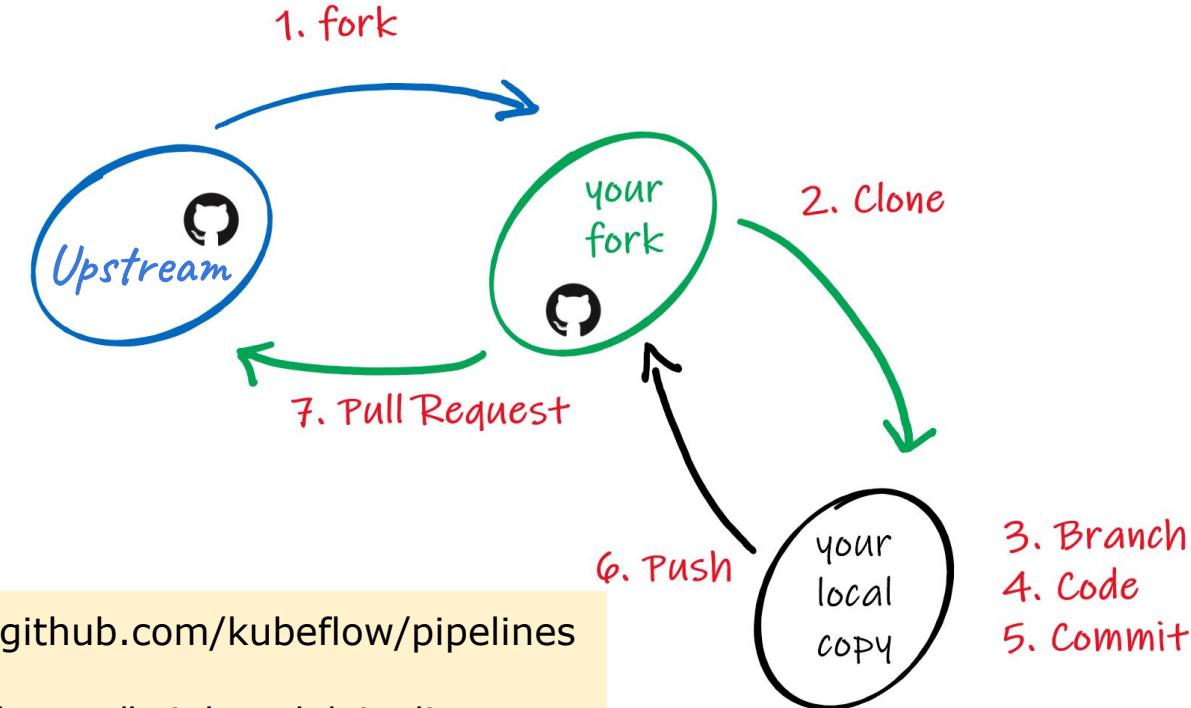
# Issue list @kubeflow/pipeline

415 Open ✓ 1,720 Closed		Author ▾	Label ▾	Projects ▾	Milestones ▾	Assignee ▾	Sort ▾
<span>⚠</span> [frontend] The sort, filter of artifact does not work	<a href="#">area/frontend</a> <a href="#">kind/bug</a>						<span>1</span>
#5565	opened 20 hours ago by capri-xiyue						
<span>⚠</span> [Doc] Update "Visualize Results in the Pipelines UI" page							<span>2</span>
#5564	opened yesterday by Bobgy						
<span>⚠</span> [backend] OutputArtifact path is None	<a href="#">area/backend</a> <a href="#">kind/bug</a>						
#5560	opened 2 days ago by wilbry						
<span>⚠</span> [feature] Configurable KFP Artifact types while storing to Minio	<a href="#">area/sdk</a> <a href="#">kind/feature</a>						<span>2</span>
#5558	opened 2 days ago by rado-narrak						
<span>⚠</span> [Multi User] Kubeflow Pipelines Missing Headers	<a href="#">area/backend</a> <a href="#">kind/bug</a>						
#5556	opened 3 days ago by PoyerRamirez						
<span>⚠</span> [feature] Allow jobs to be scheduled on AWS Fargate	<a href="#">area/backend</a> <a href="#">kind/feature</a>						<span>2</span>
#5555	opened 3 days ago by yuhuishi-convect						
<span>⚠</span> [sdk] mlpipelines-metrics path is wrong when using kfp.components.load_component_from_text	<a href="#">area/sdk</a> <a href="#">kind/bug</a>						
#5548	opened 4 days ago by santiagoolivar2017						
<span>⚠</span> pipeline_conf.set_ttl_seconds_after_finished(seconds=10) NOT WORKING	<a href="#">area/backend</a> <a href="#">kind/bug</a>						
#5545	opened 4 days ago by prasadkyp7						
<span>⚠</span> [sdk] input parameter created without default value	<a href="#">area/sdk</a> <a href="#">kind/bug</a>						<span>1</span>
#5544	opened 5 days ago by amitza						

# Git Basic



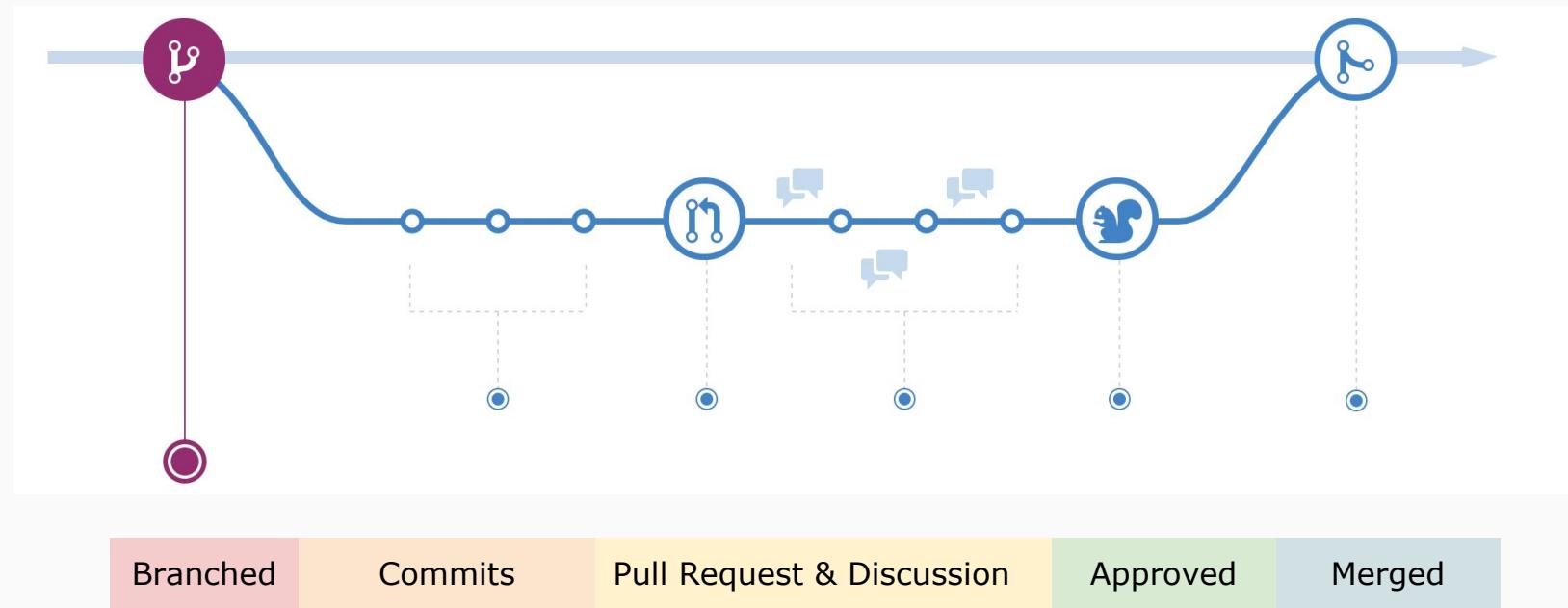
# Pull-Request flow in Github



Upstream: [github.com/kubeflow/pipelines](https://github.com/kubeflow/pipelines)

Fork: [github.com/hsinhoyeh/pipelines](https://github.com/hsinhoyeh/pipelines)

# Github Flow: branch-based workflow



# Pull-Request list @kubeflow/pipeline

81 Open ✓ 3,342 Closed

Author ▾ Label ▾ Projects ▾ Milestones ▾ Reviews ▾ Assignee ▾ Sort ▾

**chore: [Research] Add toxicity detection pipeline** ● cla: yes needs-ok-to-test size/S  
#5573 opened 1 hour ago by TheDutchDevil 1 of 1

**feat(sdk.v2): reimplement custom job support** ● cla: yes size/XL  
#5572 opened 4 hours ago by chensun 1 of 1

**Add OWNERS for v2 folder** ● cla: yes lgmt size/XS  
#5571 opened 10 hours ago by Bobgy 0 of 1

**Add zijianjyo to frontend/OWNERS** ● approved cla: yes size/XS  
#5570 opened 10 hours ago by Bobgy 0 of 1

**chore(deps): bu**  
#5569 opened 11 hours ago by TheDutchDevil

**chore(google\_c**  
size/M  
#5554 opened 3 days ago by TheDutchDevil

**feat(sdk): Add a**  
#5552 opened 3 days ago by TheDutchDevil

**chore(v2): fix in**  
size/M  
#5543 opened 5 days ago by TheDutchDevil

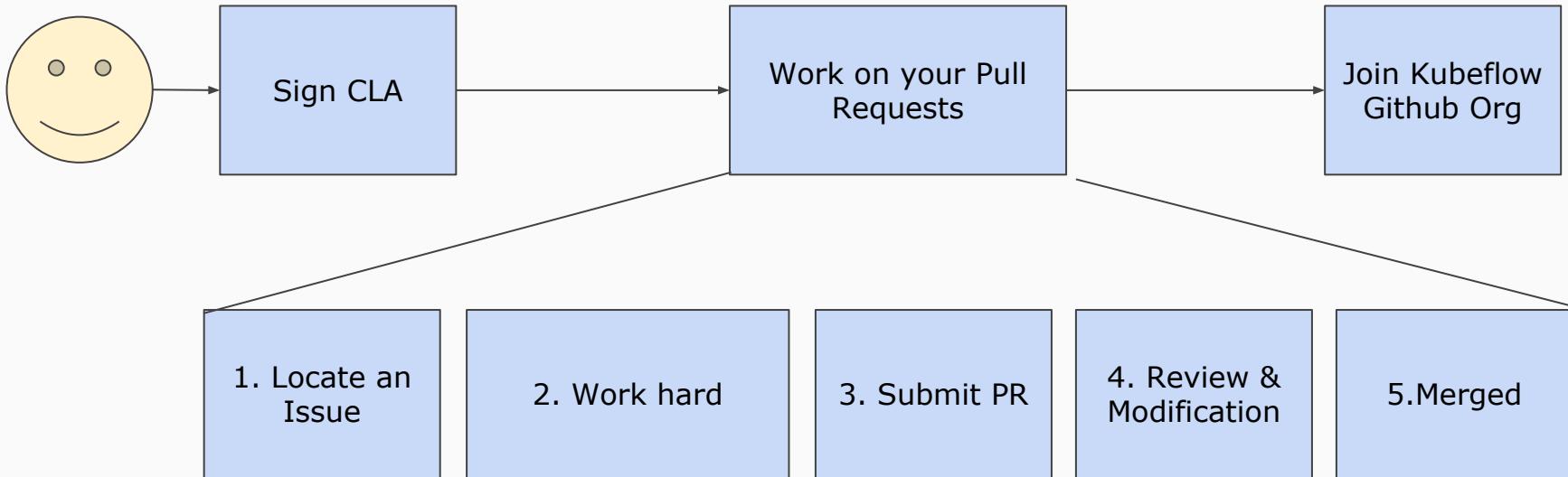
Ref: <https://www.conventionalcommits.org/en/v1.0.0/>

**<type>(<scope>): <short summary>**

- ↳ Summary in present tense. Not capitalized. No period at the end.
- ↳ Scope: common|compiler|authentication|core|
- ↳ Type: chore, docs, feat, fix, refactor, style, or test.

4 25

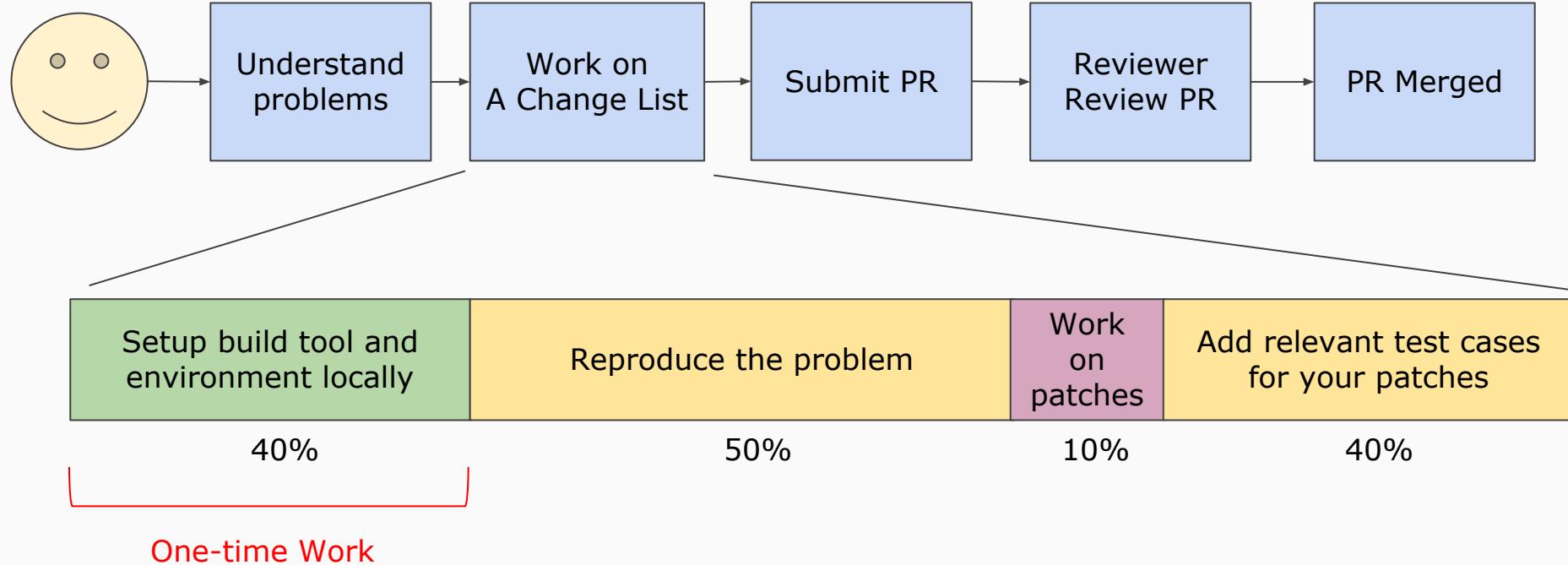
# How to be a Kubeflow Contributor?



Ref: <https://www.kubeflow.org/docs/about/contributing/>

1. Existing issues
  - a. Issues labelled with `good first issue`:  
kubeflow/pipeline
  - b. Issues labelled with `help wanted`:  
kubeflow/pipeline
2. Organic issues
  - a. Create an issue to describe issues you encountered and also provide a solution (a.k.a Pull-request) for it

# Why does the first contribution be important?



# Kubeflow Case Study

# Kubeflow/pipeline#4699 Case Study (1/2) - Existing issue

## Feature suggestion: expose the duration of each pipeline step #4699

New issue

Closed

gcampax opened this issue on Oct 31, 2020 · 1 comment



gcampax commented on Oct 31, 2020

...

It would be very useful to know how long each pipeline step took, in addition to the total duration of the pipeline. This is especially useful when trying different hyperparameters that affect performance.

This information is already available on the Argo UI, but (1) one needs to know the UI exists (it is not exposed from the Kubeflow dashboard) and (2) one needs to cross reference the pod name from the kubeflow-pipelines UI and the Argo UI. It would be nice to have the duration in one place.

5

Bobgy added this to Needs triage in KFP Runtime Triage via automation on Nov 5, 2020

Bobgy added good first issue area/frontend kind/feature help wanted labels on Nov 5, 2020



Bobgy commented on Nov 5, 2020

Collaborator

...

Contribution instructions: <https://github.com/kubeflow/pipelines/tree/master/frontend>

Run details page react component:

```
pipelines/frontend/src/pages/RunDetails.tsx  
Line 342 in e3992fa  
342 {sidepanelSelectedTab === SidePaneTab.VISUALIZATIONS &&
```

### Assignees

No one assigned

### Labels

area/frontend good first issue  
help wanted kind/feature

### Projects

KFP Runtime Triage  
Closed

### Milestone

No milestone

### Linked pull requests

Successfully merging a pull request may close this issue.  
[feat\(frontend\): Add Task Details tab whe...](#)

### Notifications

Customize

Subscribe

You're not receiving notifications from this thread.

2 participants



Bobgy moved this from Needs triage to Awaits Contributor in KFP Runtime Triage on Nov 15, 2020

## Kubeflow/pipeline#4699 Case Study (2/2) - Existing issue

feat(frontend): Add Task Details tab when clicking a node: Info includes ID, Name, Time, Status. Fixes #4699 #5023

Merged google-oss-robot merged 10 commits into `kubeflow:master` from `zijianjoy:duration` on Jan 26

Conversation 41 Commits 10 Checks 0 Files changed 5 +137 -12

**zijianjoy** commented on Jan 22 · edited

Fixes #4699

Description of your changes:  
As part of fulfilling feature request #4699, creating a Task Detail tab which allows people to view the task ID, display name, status, start time, end time, runtime duration of a node when they click on it.

Checklist:

- The title for your pull request (PR) should follow our title convention. [Learn more about the pull request title convention used in this repository](#).
- Do you want this pull request (PR) cherry-picked into the current release branch?

Screenshot:

**Success**

The screenshot shows a GitHub pull request page for a merge into the master branch. The pull request has been merged and contains 10 commits. A comment from 'zijianjoy' is visible, mentioning fix #4699 and the creation of a 'Task Detail' tab. Below the comment is a checklist with one item checked. A large red box highlights the 'lgtm' label under the 'Labels' section. To the right of the labels, the word 'LGTM?' is written in large red letters. At the bottom left, there's a diagram illustrating a pipeline task named 'Print text'. The 'Detail' tab is selected in the navigation bar, and its content displays task details like ID, name, status, and duration.

Contributor

Reviewers

- Bobby**
- capri-xiyue
- Ark-kun
- elikatsis

Assignees

- Bobby**

Labels

- Approved
- cla: yes
- lgtm**
- size/L

Projects

None yet

Milestone

No milestone

Linked issues

Successfully merging this pull request may close these issues.

Feature suggestion: expose the duration...

Notifications Customize

**LGTM?**

# Kubeflow/kubeflow#7525 Case Study - Organic issue (1/2)

[backend] health service response should be consumed by application/json not text/plain #7525

Edit

New issue

Open

hsinhyeh opened this issue 16 days ago · 2 comments



hsinhyeh commented 16 days ago

...

#### Environment

- How did you deploy Kubeflow Pipelines (KFP)? install via kubeflow manifest with version v1.4
- KFP version: v1.7.0
- KFP SDK version: v1.7.0 (golang version under backend/api/go\_http\_client)

#### Steps to reproduce

1. start pipeline backend api server by listening port 8888
2. sent http request with swagger generated code stub with the following code example

```
package main

import (
    "github.com/go-openapi/strfmt"
    "github.com/go-openapi/runtime"
    healthz_client "github.com/kubeflow/pipelines/backend/api/go_http_client/healthz_client"
    healthz_service "github.com/kubeflow/pipelines/backend/api/go_http_client/healthz_client/healthz_service"
)

func main() {
    healthServiceClient := healthz_client.New(
        client.New(
            "localhost:8888",
            healthz_client.DefaultbasePath,
            []string{"http"}),
        strfmt.NewFormats(),
    )

    clientAuthInfoWriter := runtime.ClientAuthInfoWriterFunc(func(clientRequest runtime.ClientRequest, reg)
        return nil
    ))
}
```

#### Assignees

No one assigned

#### Labels

area/backend kind/bug

#### Projects

KFP Runtime Triage  
Awaiting Contributor

#### Milestone

No milestone

#### Development

No branches or pull requests

#### Notifications

Customize

Unsubscribe

You're receiving notifications because you were mentioned.

#### 2 participants



# Kubeflow/kubeflow#7525 Case Study - Organic issue (2/2)

fix(backend): fixes healthz response by adding json content type. Fixes #7525

[Edit](#) [Code](#)

[Open](#) hsinhoyeh wants to merge 1 commit into `kubeflow:master` from `hsinhoyeh:master`

Conversation 4 Commits 1 Checks 1 Files changed 3 +117 -0

hsinhoyeh commented 14 days ago

Description of your changes:

1. explicit content type with application/json on healthz response.
2. add healthz client api under common/client for integration test
3. add integration test for healthz endpoint

Checklist:

The title for your pull request (PR) should follow our title convention. [Learn more about the pull request title convention used in this repository.](#)

fix(backend): fixes healthz response by adding json content type. Fixes 977d08e

google-oss-prov bot added the size/L label 14 days ago

google-oss-prov bot commented 14 days ago

Hi @hsinhoyeh. Thanks for your PR.

I'm waiting for a `kubeflow` member to verify that this patch is reasonable to test. If it is, they should reply with `/ok-to-test` on its own line. Until that is done, I will not automatically test new commits in this PR, but the usual testing commands by org members will still work. Regular contributors should [join the org](#) to skip this step.

Once the patch is verified, the new status will be reflected by the `ok-to-test` label.

I understand the commands that are listed [here](#).

Details

Reviewers

- zijianjoy
- chensun
- Linchin

Assignees

No one assigned

Labels

- needs-ok-to-test
- size/L

Projects

None yet

Milestone

No milestone

Development

Successfully merging this pull request may close these issues.

None yet

Notifications

Customize

Unsubscribe

You're receiving notifications because you were mentioned.

# Kubeflow/pipeline#5493 Case Study: kind/question

default resource overhead for each individual profile #5493

Open hsinhyeh opened this issue 15 days ago · 6 comments

 hsinhyeh commented 15 days ago

**Question:**  
Not sure whether this question has been discussed before, but I am curious about the resource overhead for each individual profile.  
In the current configuration (in v1.2, v1.3), when creating one single profile, there are two running pods are created by default:  
  
`ml-pipeline-ui-artifact-c8fbcb8f9-k88bz  
ml-pipeline-visualizationserver-6b78c9646-s7q4c`  
  
each one takes additional istio sidecar container (cost 100m cpu/ 128Mi memory).  
so in total cpu requested  $100m * 2 + 128m * 2 = 0.5$  CPU unit, memory requested  $128 * 2 + 128 * 2 = 512Mi$ .  
(even the profile has been created but no further actions were conducted.)  
  
so let say today we have to create 100 profiles (one for each individual participant) but only 20% of them are heavy users (80-20 rule), and our system wasted  $100 * 80\% * 0.5$  CPU unit (and 512Mi + 100 \* 80% memory).  
  
This seems to be more like a resource leaking, not the design concept. any thoughts? or documents? thanks

transferred from [kubeflow/kubeflow#5832](#)

 zjianjoy commented 8 days ago

Hello @hsinhyeh , you can disable the pipeline pods by setting the following label from your profile's namespace labels: `pipelines.kubeflow.org/enabled: false` .  
  
You can do this for all the profiles that don't require Kubeflow Pipelines. For future roadmap, we plan to remove visualization server too.  
  
Does this solve your issue?

1

Edit New issue

**Assignees**  
No one assigned

**Labels**  
needs more info

**Projects**  
KFP Runtime Triage  
Needs More Info

**Milestone**  
No milestone

**Linked pull requests**  
Successfully merging a pull request may close this issue.  
None yet

**Notifications** Customize Unsubscribe  
You're receiving notifications because you were mentioned.

**3 participants**  


# Kubeflow/katib#1484 Case Study: kind/discussion

Open Katib Roadmap 2021 #1484  
andreyvelich opened this issue on Mar 19 · 7 comments · May be fixed by #1524

We should discuss our Katib Roadmap for 2021.

Some points from my side:

### Katib enhancements

- Finish AWS CI/CD migration for the release process.
- New Katib UI implementation: <https://github.com/kubeflow/katib/projects/1>.
- Support [multiple versions](#) in Katib CRDs.
- Support various parameter distribution: #1207.
- Finish validation for the algorithms: #1126.

### Advanced AutoML features

- Refactor Hyperband: #1389.
- Support ASHA, I believe it is very popular.
- Separate CRD for NAS and HP ?
- Support Population Based Training: #1382.
- Support AutomC: #460.
- Support Feature Engineering: #475.

Please give your feedback and what do you want to see in Katib 2021.  
cc @gaocegege @johnugeorge @PatrickXYS @terrytangyuan @kimwnasptd @yanniszark @jbottum @c-bata @RFMVASCONCELOS @knkski @anneum @thomas-mueller

Let's finalise our Roadmap on the upcoming WG meetings (24th of March and 7th of April).

5 6

 andreyvelich added the kind/discussion label on Mar 19

 PatrickXYS commented on Mar 19 Member ...

No one assigned

Labels  
[kind/discussion](#)

Projects  
None yet

Milestone  
No milestone

Linked pull requests  
Successfully merging a pull request may close this issue.

 Add Katib 2021 ROADMAP

Notifications      Customize  


You're receiving notifications because you're subscribed to this thread.

4 participants  


# Kserve/Kserve#2355 Case Study: kind/feature

## Add an example of Lime text explainer from AIX360 toolkit #2355

Merged

kserve-oss-bot merged 9 commits into kserve:master from C1berwiz:master on Oct 16, 2022

Conversation 26

Commits 9

Checks 43

Files changed 7

+236 -6



C1berwiz commented on Jul 26, 2022 • edited by yuzisun

Contributor



...

### What this PR does / why we need it:

This PR added an example of LIME text explainer of AIX360 toolkit

Since there are 3 modes of LIME explainer (image, text and tabular ), but currently there is only a sample of image (MNIST)

We suppose that adding a text example would help others better understanding LIME explainer

### Type of changes

Please delete options that are not relevant.

New feature (non-breaking change which adds functionality)

### Feature/Issue validation/testing:

Please describe the tests that you ran to verify your changes and relevant result summary. Provide instructions so it can be reproduced.

Please also list any relevant details for your test configuration.

Test A

- like the mnist sample of aix and mentioned in README.md , this sample should execute just like mnist sample
- The first step is to [determine the ingress IP and ports](#) and set INGRESS\_HOST and INGRESS\_PORT

```
MODEL_NAME=aix-explainer
SERVICE_HOSTNAME=$(kubectl get inferenceservice ${MODEL_NAME} -o jsonpath='{.status.url}' | cut -d "/" -f 3)
python query_explain.py http://${INGRESS_HOST}:${INGRESS_PORT}/v1/models/${MODEL_NAME}:explain ${SERVICE_HOSTNAME}
```

### Reviewers

yuzisun



theofpa



yuzliu



drewbutlerbb4



Tomcli



Still in progress? Learn about draft PRs



### Assignees

yuzisun

### Labels

approved lgtm

### Projects

None yet

### Milestone

No milestone

# Other Case Study

# nats-io/nats.go#22 Case Study

 **Closed** goroutine leaking when calls BindRecvChan #22  
hsinhyeh opened this issue on Jul 26, 2014 · 1 comment

in the underlying implementation, nats defines a callback (i.e. cb) and then call the Subscribe from a conn  
<https://github.com/apcera/nats/blob/master/netchan.go#L81>

```
_ , err := c.Conn.subscribe(subject, queue, cb)
```

in the conn.subscribe, nats spawns a go routine to send the message:  
<https://github.com/apcera/nats/blob/master/nats.go#L1198>

```
go nc.deliverMsgs(sub.mch)
```

and that go routine will be terminated "only if sub.mch is closed", which is called via conn.Unsubscribe().  
<https://github.com/apcera/nats/blob/master/nats.go#L1244>

```
func (nc *Conn) unsubscribe(sub *Subscription, max int) error {
```

The bug actually happened, when we look BindRecvChan again: even the caller has closed the channel, the go routine who serves deliverMsg will never be destroyed. We actually have an additional zombie go routine when we call BindRecvChan every time.

hsinhyeh referenced this issue in hsinhyeh/nats on Jul 26, 2014

derekcollison added the **bug** label on Jul 26, 2014

derekcollison self-assigned this on Jul 26, 2014

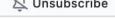
derekcollison commented on Jul 26, 2014

The proposed solution could affect ordering of messages due to the read. But great bug report. Will fix. Thanks.

Projects  
None yet

Milestone  
No milestone

Linked pull requests  
Successfully merging a pull request may close this issue.  
None yet

Notifications  
Customize  
 Unsubscribe  
You're receiving notifications because you were assigned.

2 participants  


# googleapis/google-cloud-go#170 Case Study (1/2)

[bigtable] bttest needs UpdateColumnFamily support #170  
hsinohoyeh opened this issue on Sep 15, 2015 · 3 comments

diff --git a/google.golang.org/cloud/bigtable/bttest/inmem.go b/google.golang.org/cloud/bigtable/bttest/inmem.go  
index 5c4bc09..f4dbe03 100644  
--- a/google.golang.org/cloud/bigtable/bttest/inmem.go  
+++ b/google.golang.org/cloud/bigtable/bttest/inmem.go  
+func (s \*server) UpdateColumnFamily(ctx context.Context, req \*bttdpb.ColumnFamily) (\*bttdpb.ColumnFamily, error)  
+ return nil, errors.New("no op")  
+}

we need it because we call SetGCPolicy to configure the maximum versions stored in a cell

```
adminClient.SetGCPolicy( ctx, tableName, columnFamily, bgt.MaxVersionsPolicy(1))
```

and in the underlying implementation, it will first call GetTable and then update its column family.

```
func (ac *AdminClient) SetGCPolicy(ctx context.Context, table, family string, policy GCPolicy) error {  
    prefix := ac.clusterPrefix()  
    tbl, err := ac.tClient.GetTable(ctx, &bttspb.GetTableRequest{  
        Name: prefix + "/tables/" + table,  
    })  
    if err != nil {  
        return err  
    }  
    fam, ok := tbl.ColumnFamilies[family]  
    if !ok {  
        return fmt.Errorf("unknown column family %q", family)  
    }  
    fam.GcRule = policy.proto()  
    _, err = ac.tClient.UpdateColumnFamily(ctx, fam)  
    return err  
}
```

<https://github.com/GoogleCloudPlatform/gcloud-golang/blob/master/bigtable/admin.go#L161>

hsinohoyeh commented on Sep 16, 2015

hi,  
I think I have fixed this (#170) and #166. you can see more in below.  
you can purify my CL and push to the repo or I can create a RP if you think the CL is ok

api:bigtable triage me

Projects  
None yet

Milestone  
No milestone

Linked pull requests  
Successfully merging a pull request may close this issue.  
None yet

Notifications Unsubscribe  
You're receiving notifications because you authored the thread.

4 participants

# googleapis/google-cloud-go#170 Case Study (2/2)

Gerrit CHANGES DOCUMENTATION BROWSE

Merged 3471 bigtable/bttest: support GetTable/UpdateColumnFamily

Change Info SHOW ALL ▾

Submitted Sep 23, 2015  
Owner Hsin-Ho Yeh  
Uploader David Symonds  
Committer Gerrit Code Review  
Reviewers David Symonds  
Repo | Branch gocloud | master  
Topic

bigtable/bttest: support GetTable/UpdateColumnFamily  
Fixes #166.  
Fixes #170.  
Change-Id: [tc2b3698583923042253d591b137c7192605ecef0](#)

Comments 25 resolved

Submit requirements

✓ Code-Review +2 David Symonds

Files	Comments	Findings	DOWNLOAD	INCLUDED IN	EXPAND ALL
Base → Patchset 4   <a href="#">bccae9</a>					
File					
Commit message					
<a href="#">bigtable/bttest/inmem.go</a>					

Change Log EXPAND ALL

Hsin-Ho Yeh	Uploaded patch set 1.	Patchset 1   Sept 19, 2015 16:32
David Symonds	16 comments Thanks for doing this. Please check the CLA details first. If you're working for a company then you might need to get them to approve your contribution.	Patchset 1   Sept 22, 2015 13:10
Hsin-Ho Yeh	Uploaded patch set 2.	Patchset 2   Sept 22, 2015 23:05
Hsin-Ho Yeh	16 comments the updated patch address the comments.	Patchset 2   Sept 22, 2015 23:19
David Symonds	10 comments Nearly there.	Patchset 2   Sept 23, 2015 07:12

# torvalds/linux#661 Case Study - Not all CLs are valuable

**Open** Fixed typo: extra space. #661  
jacob-macleod wants to merge 1 commit into `torvalds:master` from `jacob-macleod:master`

jacob-macleod commented on Mar 8, 2019

There was a double space.

10 2 13 1 6

-o Fixed typo: extra space. ... Verified ee9ca9c

KernelPRBot commented on Mar 8, 2019

Hi @jacob-and-nathan!

Thanks for your contribution to the Linux kernel!

Linux kernel development happens on mailing lists, rather than on GitHub – this GitHub repository is a read-only mirror that isn't used for accepting contributions. So that your change can become part of Linux, please email it to us as a patch.

Sending patches isn't quite as simple as sending a pull request, but fortunately it is a well documented process.

Here's what to do:

- Format your contribution according to kernel requirements
- Decide who to send your contribution to
- Set up your system to send your contribution as an email
- Send your contribution and wait for feedback

### How do I format my contribution?

The Linux kernel community is notoriously picky about how contributions are formatted and sent. Fortunately, they have documented their expectations.

Firstly, all contributions need to be formatted as patches. A patch is a plain text document showing the change you want to make to the code, and documenting why it is a good idea.

You can create patches with `git format-patch`.

Secondly, patches need 'commit messages' which is the human-friendly documentation explaining what the change is

**Reviewers**  
ssbarnea ✓  
StackNeverFlow ✓

**Assignees**  
No one assigned

**Labels**  
None yet

**Projects**  
None yet

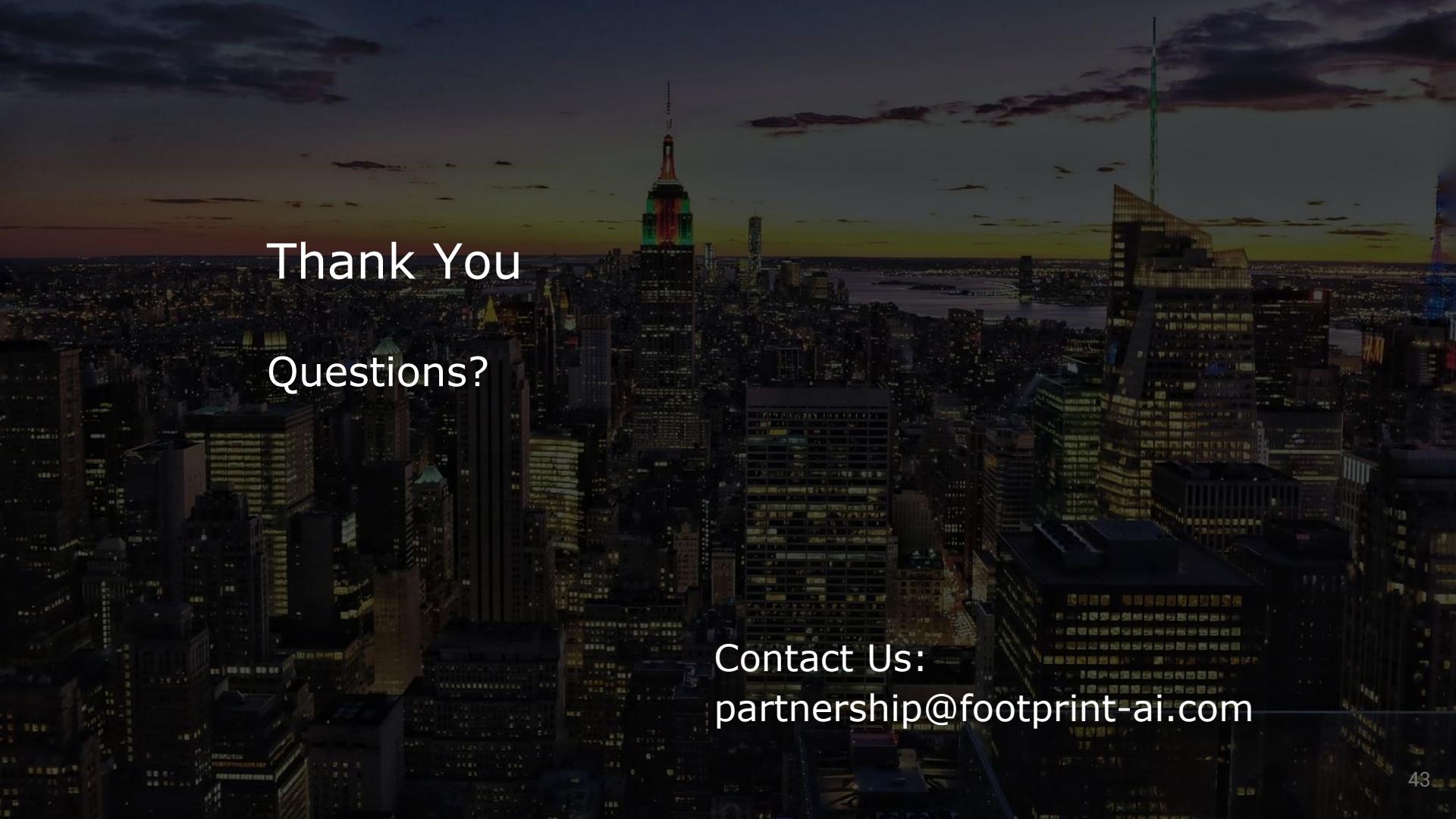
**Milestone**  
No milestone

**Notifications** Customize Subscribe  
You're not receiving notifications from this thread.

**5 participants**

# Conclusion

- Open source is a public artifacts that help you grow a reputation and a career.
- The work to contribute to Open Source could take days or weeks, but the rewards could keep years.
- Let's make changes to the software world now, even steps are small :)



Thank You

Questions?

Contact Us:  
[partnership@footprint-ai.com](mailto:partnership@footprint-ai.com)