

# Final Project Report

## Digital Business

2022 CS-5270 Cloud Native Development

**Group10** : 游智鈞/高承翰/詹銓傑/官學勤/廖家鴻  
2022/6/8

# Outline

- Introduction
- Basic
  - Application Architecture
  - Code Optimization
  - Testing Coverage
  - CI/CD
- Option
  - App Monitor
  - Database
  - Code Contribution
- Discussion
- Conclusion

# Introduction

(廖家鴻)

# Requirement Coverage

Type	Category	Remark
Basic	Upload codes to Github	
	Deploy to GKE	
	Digital Business: Internet volume changes	
	Application Architecture	
	Code Optimization	
	Testing Coverage	
	CI / CD	
Option	Application Monitor	
	Use Database	
	Code Contribution	
	Ingress	

# Upload codes to Github

CHKao777 / tsmc-project Public

Notifications Fork 3

<> Code Issues Pull requests Actions Projects Wiki Security Insights

master 3 branches 0 tags Go to file Code

AaronKFC add k8s CD yaml file 1 0c6ba72 15 hours ago 56 commits

.github/workflows	Update python-package-conda.yml	2 days ago
crawler	add test	2 days ago
figures	add files	15 hours ago
frontend	add rq deployment (#6)	4 days ago
k8s	add k8s CD yaml file	15 hours ago
nodeserver	revise index.js	3 days ago
tests	add test	2 days ago
yaml	ingress update	12 hours ago
.DS_Store	add homepage and can render to each other page	4 days ago
.gitignore	add integration test	2 days ago
Dockerfile	add nodeserver	7 days ago
README.md	add image	15 hours ago
deploy.yaml	redis throughput test	2 days ago
docker-compose.yml	redis throughput test	2 days ago
environment.yml	add test	2 days ago

About

No description, website, or to

Readme

0 stars

2 watching

3 forks

Releases

No releases published

Packages

No packages published

Contributors 5

jayzhan211 Jay Zhan

CHKao777 CHKao

WesleyGuan gxq

strupfrmnth Thomas Yu

Github link:

<https://github.com/CHKao777/tsmc-project>

# Deploy to GKE

Pods

Services

Ingress

Database

Google Cloud Platform

nycu-lab-10

Search

Products, resources, docs (/)

Kubernetes Engine

Clusters

Workloads

Services & Ingress

Marketplace

Release Notes

Clusters

EDIT

DELETE

DEPLOY

CONNECT

DUPLICATE

tsmc-cluster

DETAILS

STORAGE

LOGS

Cluster basics

Name	tsmc-cluster	
Location type	Regional	
Region	asia-east1	

CLOUD SHELL

Terminal

(nycu-lab-10) x (nycu-lab-10) x +

```
climnehcc234@cloudshell:~ (nycu-lab-10) $ kubectl get pod
W0607 17:49:25.663173 17645 gcp.go:120] WARNING: the gcp auth plugin is deprecated in v1.22+, unavailable in v1.25+; use gcloud instead.
To learn more, consult https://cloud.google.com/blog/products/containers-kubernetes/kubectl-auth-changes-in-gke
```

NAME	READY	STATUS	RESTARTS	AGE
crawler-27574985-fdg74	0/1	Completed	0	34h
frontend-58699f8dc4-8fvwr	1/1	Running	0	2d5h
frontend-58699f8dc4-pbj92	1/1	Running	0	2d5h
frontend-58699f8dc4-vz84j	1/1	Running	0	2d5h
mongo-56cf6b45fc-tnpd6	1/1	Running	0	2d7h
nodeserver-7997576f6d-p287r	1/1	Running	0	32m
redis-9b88c854c-mdkrq	1/1	Running	0	2d7h
worker-7f47c8d444-4h5qd	1/1	Running	0	2d7h
worker-7f47c8d444-7vgr6	1/1	Running	0	2d7h
worker-7f47c8d444-gdhvj	1/1	Running	0	2d7h
worker-7f47c8d444-x2bc6	1/1	Running	0	2d7h

```
climnehcc234@cloudshell:~ (nycu-lab-10) $ kubectl get service
W0607 17:49:31.995676 17663 gcp.go:120] WARNING: the gcp auth plugin is deprecated in v1.22+, unavailable in v1.25+; use gcloud instead.
To learn more, consult https://cloud.google.com/blog/products/containers-kubernetes/kubectl-auth-changes-in-gke
```

NAME	TYPE	CLUSTER-IP	EXTERNAL-IP	PORT(S)	AGE
frontend	NodePort	10.28.130.108	<none>	5000:31577/TCP	2d7h
mongodb	ClusterIP	10.28.129.120	<none>	27017/TCP	2d7h
nodeserver-service	NodePort	10.28.129.70	<none>	3000:30036/TCP	78m
redis	ClusterIP	10.28.130.122	<none>	6379/TCP	2d7h

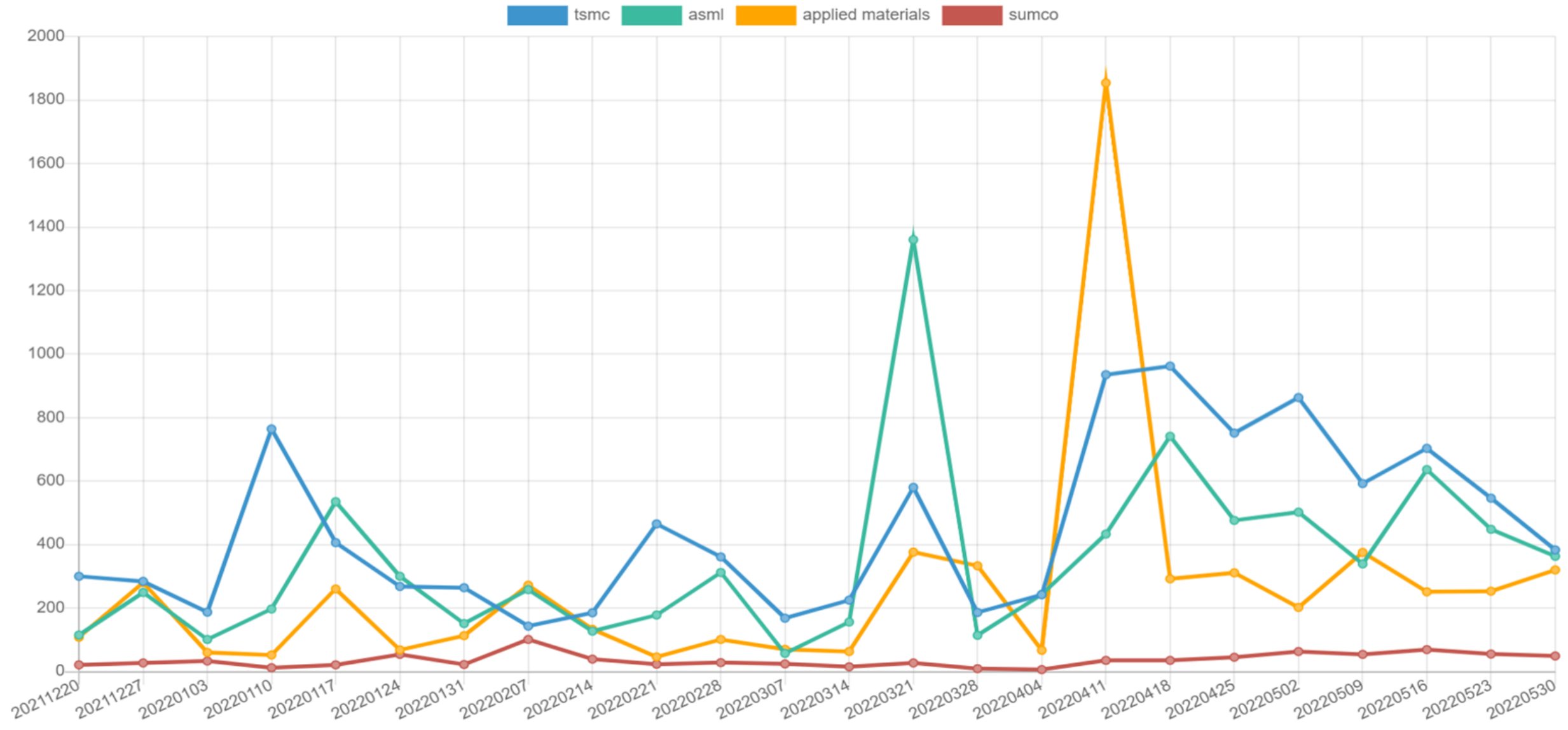
```
climnehcc234@cloudshell:~ (nycu-lab-10) $ kubectl get ingress
W0607 17:49:36.607262 17681 gcp.go:120] WARNING: the gcp auth plugin is deprecated in v1.22+, unavailable in v1.25+; use gcloud instead.
To learn more, consult https://cloud.google.com/blog/products/containers-kubernetes/kubectl-auth-changes-in-gke
```

NAME	CLASS	HOSTS	ADDRESS	PORTS	AGE
nodeserver-ingress	<none>	*	34.120.17.241	80	4d1h

```
climnehcc234@cloudshell:~ (nycu-lab-10) $ kubectl get PersistentVolumeClaim
W0607 17:51:54.358532 17830 gcp.go:120] WARNING: the gcp auth plugin is deprecated in v1.22+, unavailable in v1.25+; use gcloud instead.
To learn more, consult https://cloud.google.com/blog/products/containers-kubernetes/kubectl-auth-changes-in-gke
```

NAME	STATUS	VOLUME	CAPACITY	ACCESS MODES	STORAGECLASS	AGE
mongo-pvc	Bound	pvc-a7c0ced9-5c7d-4634-bb10-d374c397cfea	1Gi	RWO	standard-rwo	2d7h

# Digital Business: Internet Volume (word count)



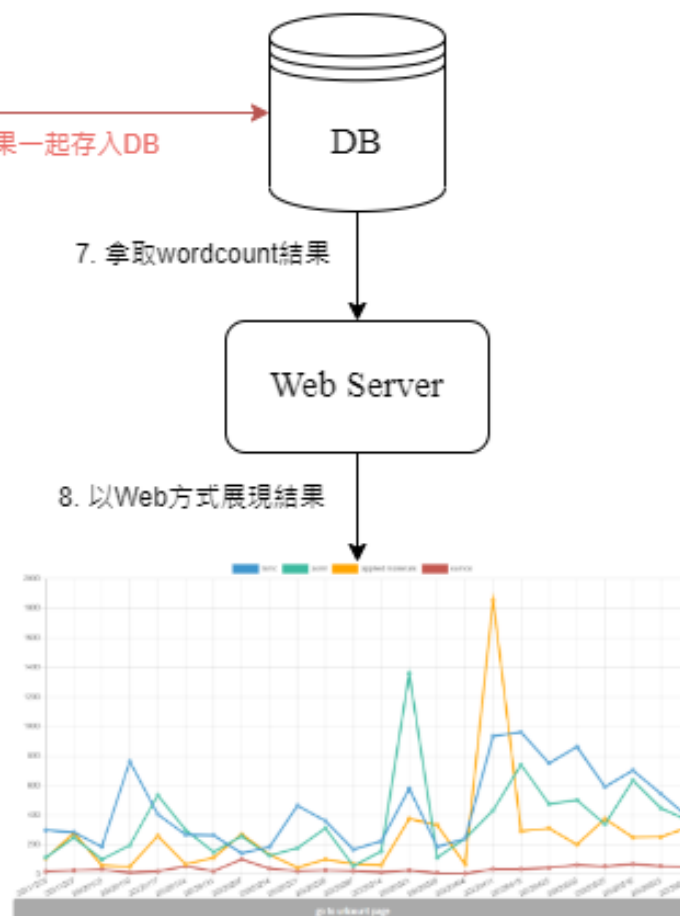
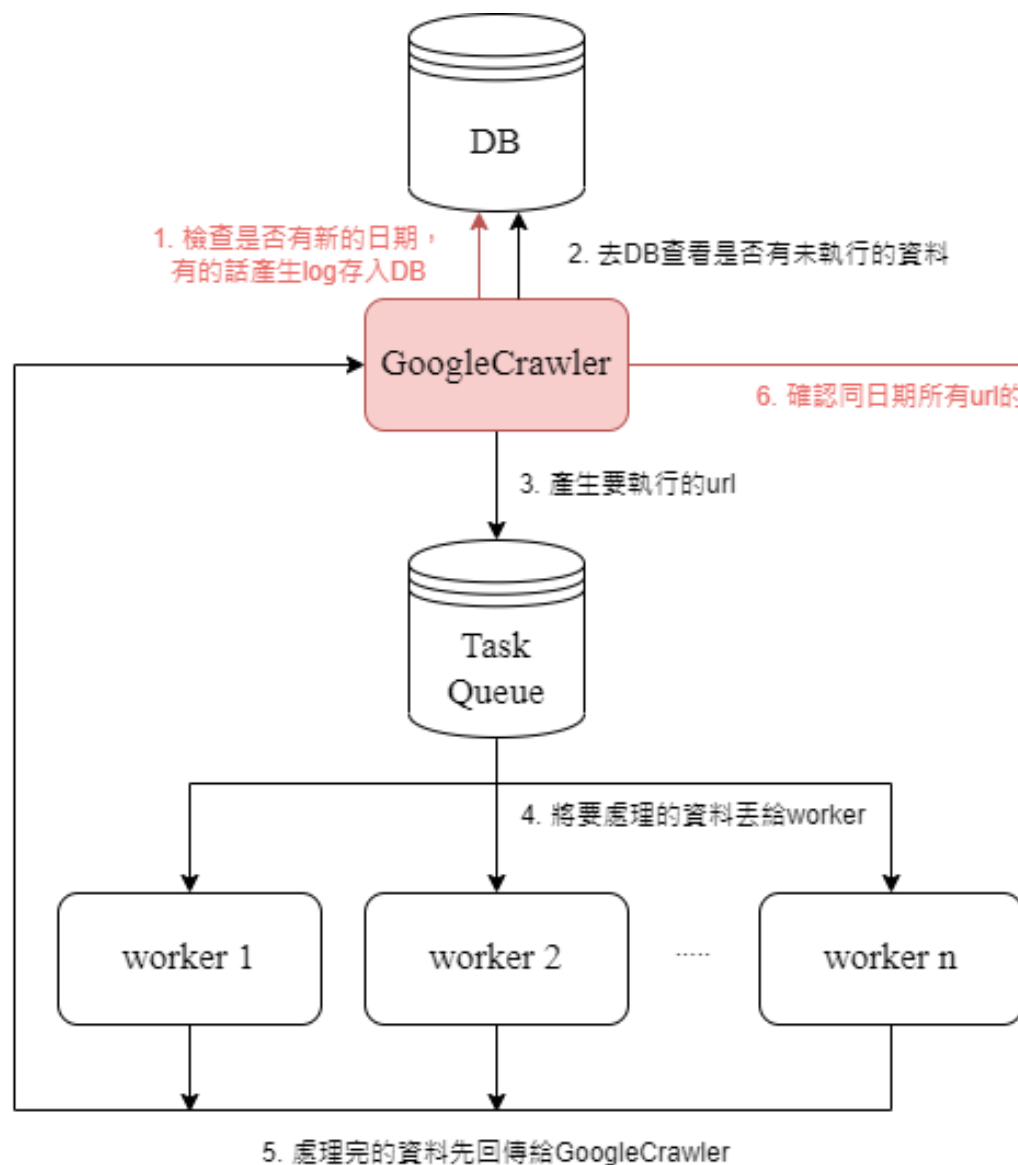
Service link: <http://34.120.17.241/wordcount>

# Application Architecture

(游智鈞)

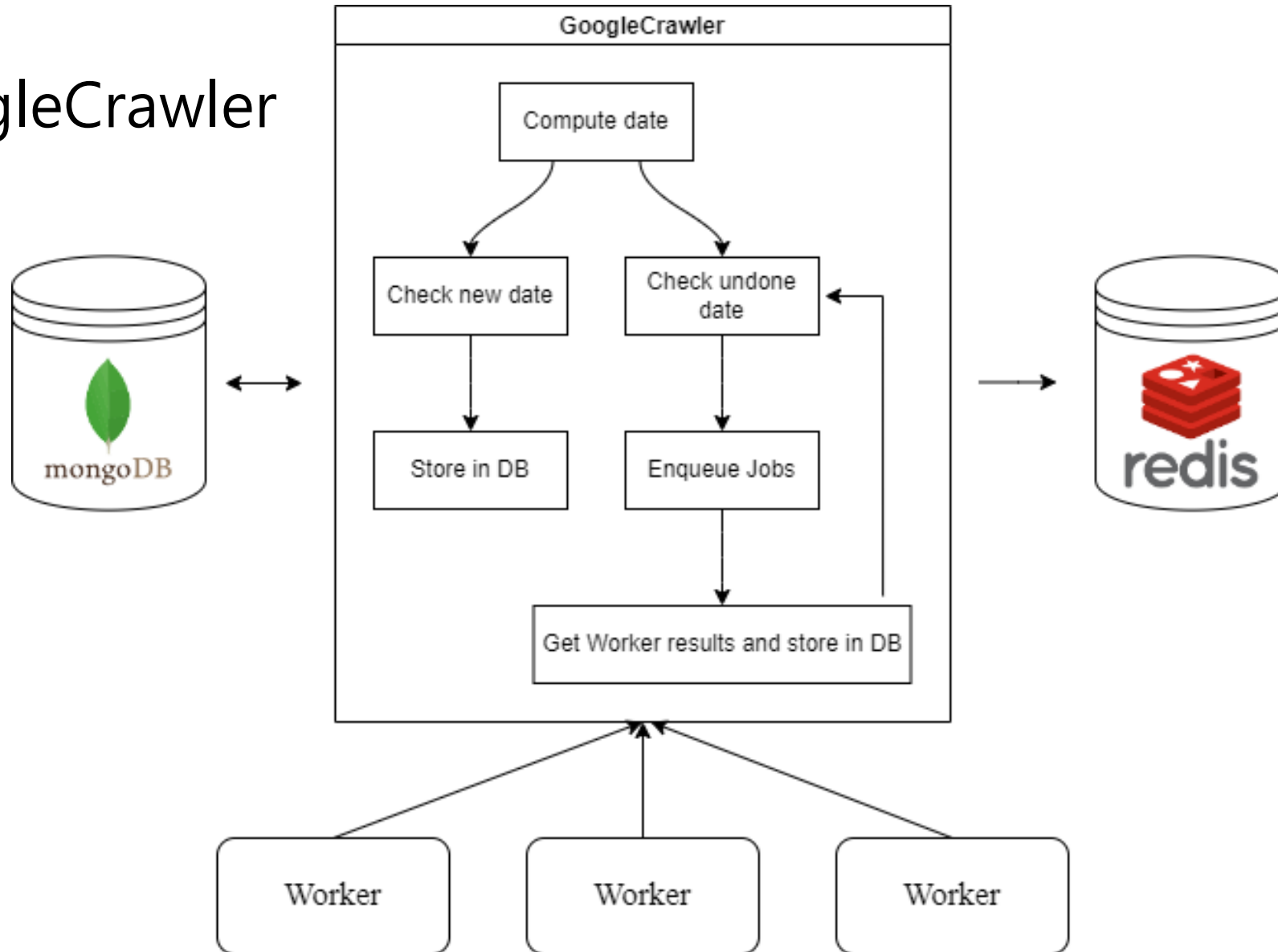


# Application Architecture



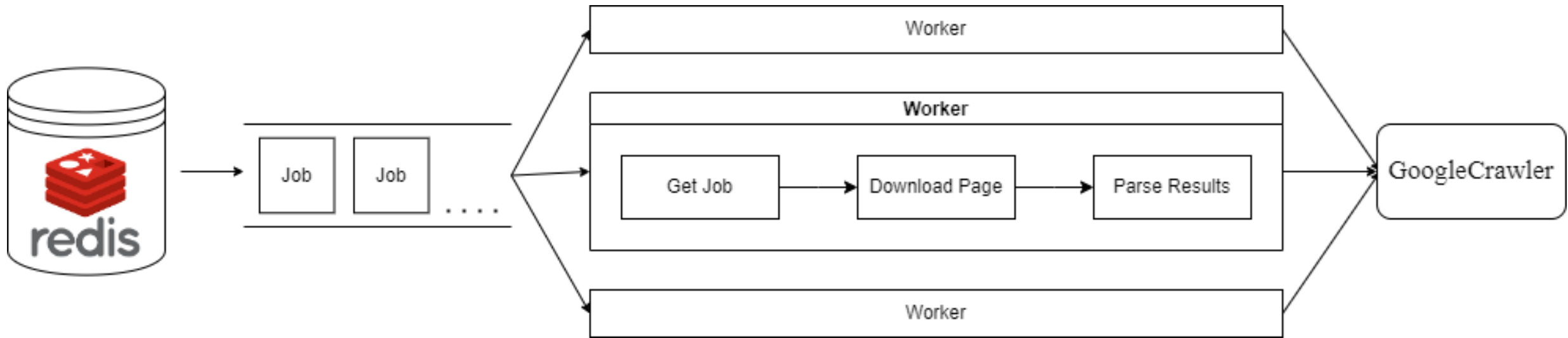
# Application Architecture

- GoogleCrawler



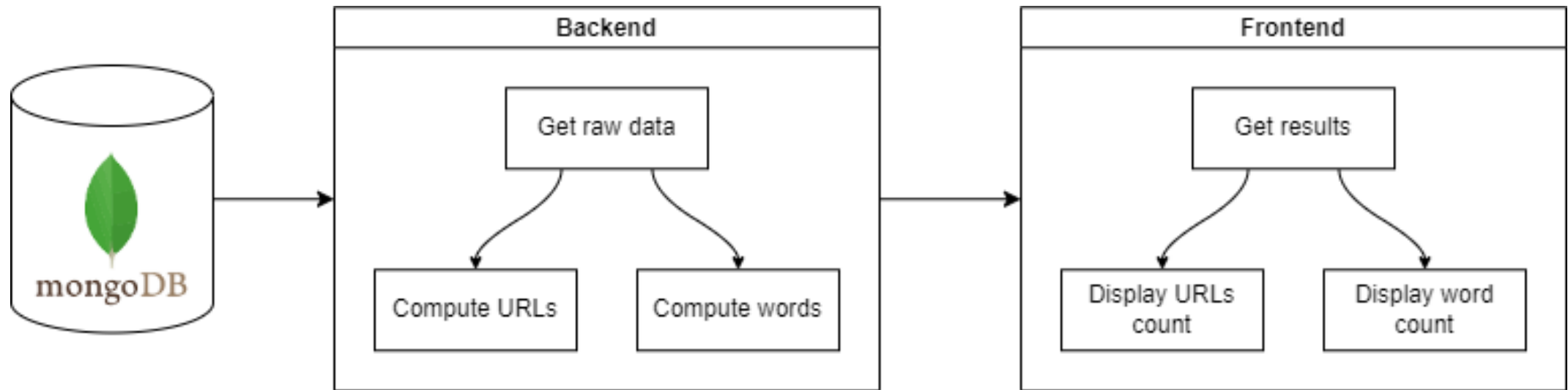
# Application Architecture

- Crawler



# Application Architecture

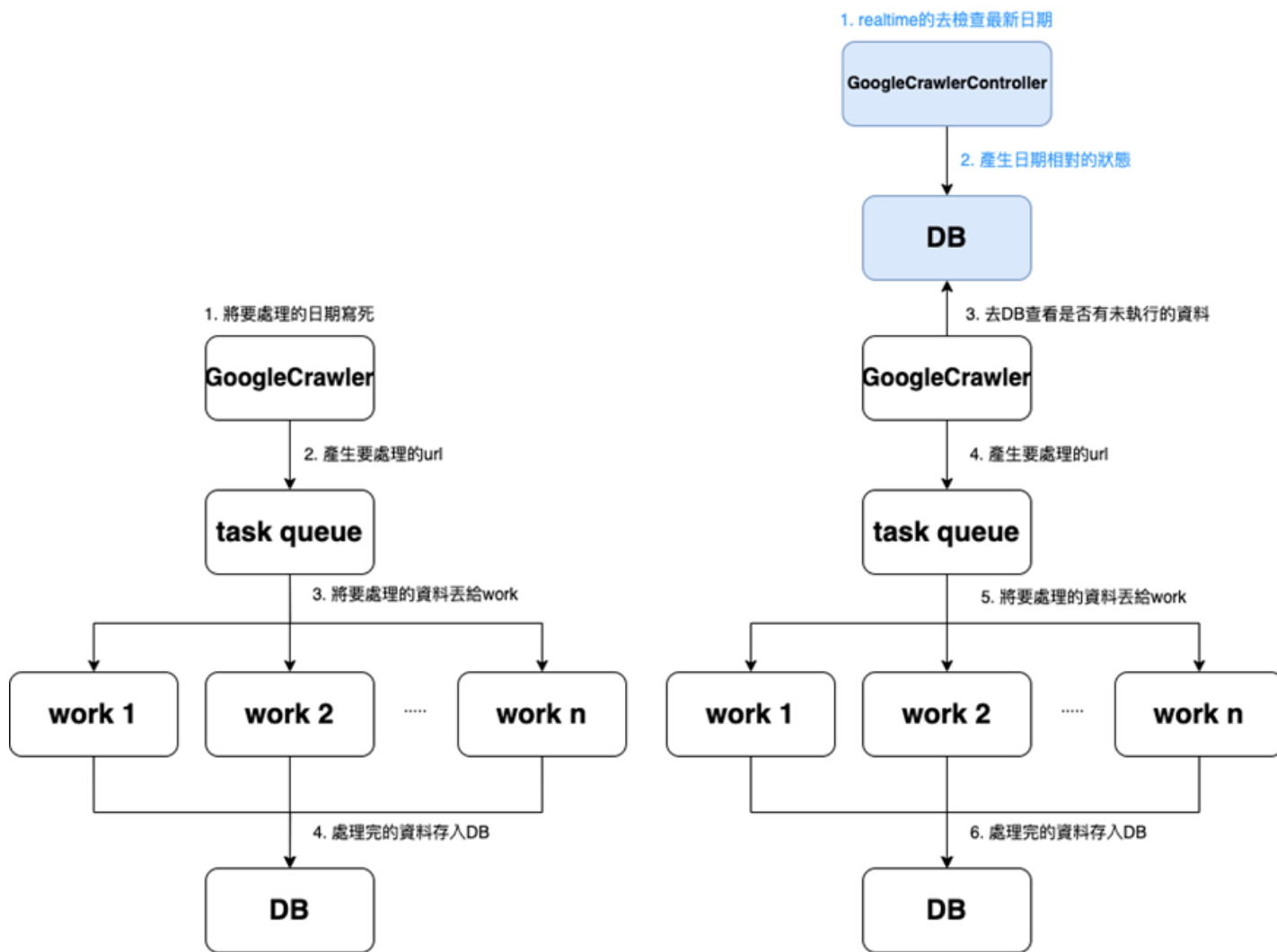
- Web Server



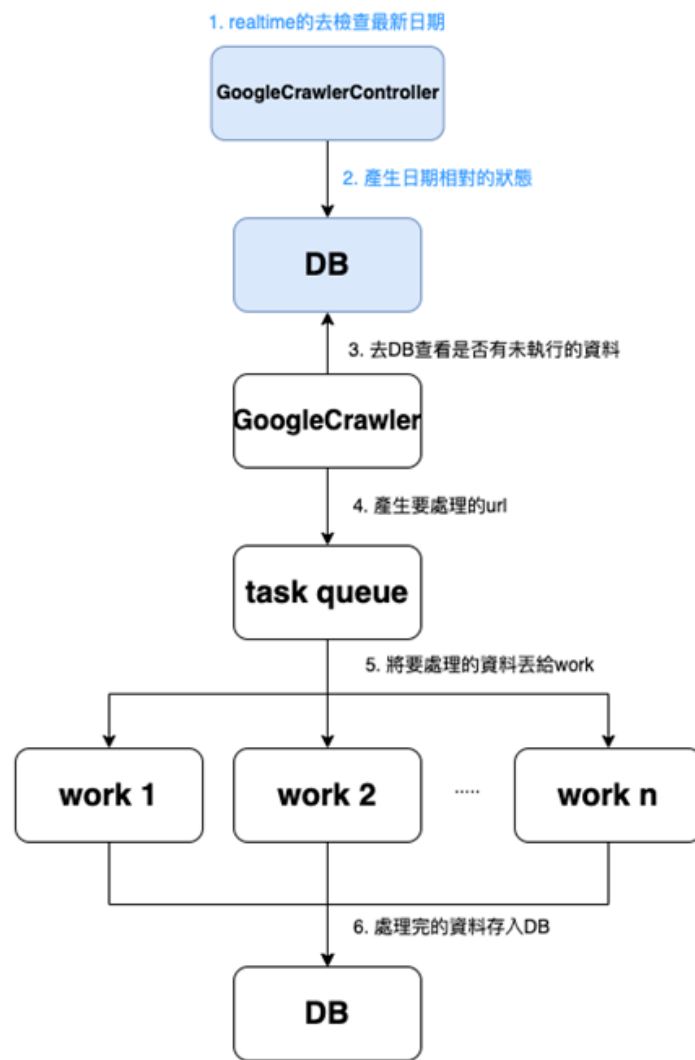
# Code Optimization

(官學勤)

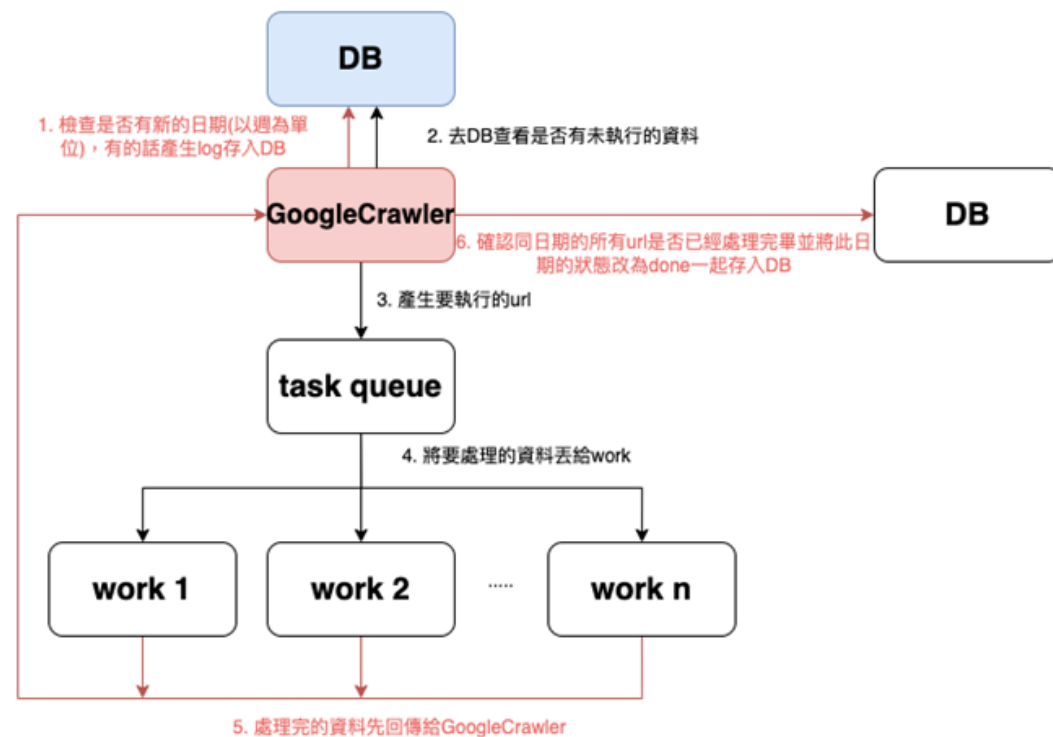
# Code Optimization



5/13 local k8s version



5/25 local k8s version



6/5 GKE version

# CronJob

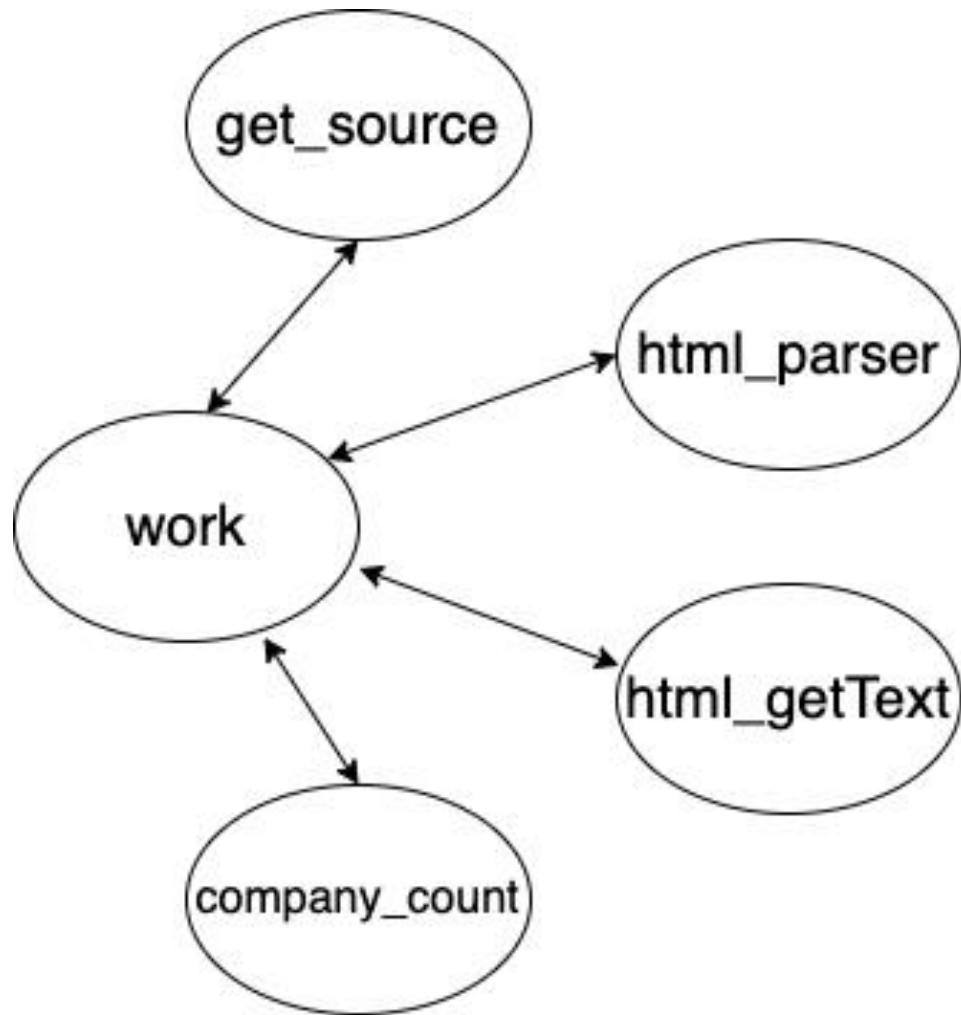
## Cron schedule syntax

```
# |----- minute (0 - 59)
# |----- hour (0 - 23)
# |----- day of the month (1 - 31)
# |----- month (1 - 12)
# |----- day of the week (0 - 6)
# |----- 7 :
# |----- OR
# * * * * *
```

```
kind: CronJob
metadata:
  name: crawler
spec:
  schedule: "0 0 * * 1"
  concurrencyPolicy: Forbid
  successfulJobsHistoryLimit: 1
  failedJobsHistoryLimit: 1
  jobTemplate:
    spec:
      template:
        spec:
          containers:
            - name: crawler
              image: chkao777/tsmc-project:crawler
              imagePullPolicy: Always
          restartPolicy: OnFailure
```

At 00:00 on Monday.

# Simple Responsibility Principle



```
def work(url, timestamp):  
    response = get_source(url)  
    if response is None:  
        return None  
    soup = html_parser(response.text)  
    if soup is None:  
        return None  
    original_text = html_getText(soup)  
    return company_count(original_text, timestamp)
```

```
def get_source(url): ...
```

```
def html_parser(htmlText): ...
```

```
def html_getText(soup): ...
```

```
def company_count(text, timestamp): ...
```



# Testing Coverage

(詹銓傑)

# Unit test

- Test Worker
  - Company count
  - Get HTML Text (Only get word in paragraph)
- Test Crawler
  - Add new date
  - Get first weekday (Monday)
  - Test with Mongodb (pytest-mongodb)

```
===== test session starts =====
platform linux -- Python 3.10.4, pytest-6.2.5, py-1.11.0, pluggy-1.0.0
rootdir: /home/jayzhan/tsmc-project/tests, configfile: pytest.ini
plugins: integration-0.2.2, cov-3.0.0, mongodb-2.2.0, datafiles-2.0.1
collected 7 items

tests/test_crawler.py .... [ 57%]
tests/test_worker.py .. [ 85%]
tests/integrations/test_worker_integreation.py . [100%]

===== 7 passed in 1.14s =====
```

# Integration test

- Test worker
- URL: '<https://www.taipeitimes.com/News/biz/archives/2022/01/20/2003771688>'
- tsmc: 4, asml: 7

```
===== test session starts =====
platform linux -- Python 3.10.4, pytest-6.2.5, py-1.11.0, pluggy-1.0.0
rootdir: /home/jayzhan/tsmc-project/tests, configfile: pytest.ini
plugins: integration-0.2.2, cov-3.0.0, mongoddb-2.2.0, datafiles-2.0.1
collected 7 items

tests/test_crawler.py .... [ 57%]
tests/test_worker.py .. [ 85%]
tests/integrations/test_worker_integreation.py . [100%]

===== 7 passed in 1.14s =====
```

# Testing Coverage

- GoogleCrawler.py Coverage 26%
- GoogleCrawlerTest.py Coverage 0% (Not used)
- worker.py Coverage 69%
- Total Coverage 32%

```
tests/test_crawler.py .... [ 57%]
tests/test_worker.py .. [ 85%]
tests/integrations/test_worker_integreation.py . [100%]

----- generated xml file: /home/jayzhan/tsmc-project/pytest.xml -----

----- coverage: platform linux, python 3.10.4-final-0 -----
Name                               Stmts  Miss  Cover  Missing
-----
crawler/GoogleCrawler.py           144    106    26%  20-37, 40-45, 49-52, 56-62, 66-75, 81-94, 103-111, 116-150, 153-171, 200-222
crawler/GoogleCrawlerTest.py        49     49     0%  1-76
crawler/worker.py                   59     18    69%  20-27, 30-35, 40-41, 79-80
-----
TOTAL                               253    173    32%

1 file skipped due to complete coverage.

===== 7 passed in 1.46s =====
```

CI / CD

# Github Action

CHKao777 / tsmc-project Public

Notifications

Fork 3

Star 0

<> Code Issues Pull requests **Actions** Projects Wiki Security Insights

## Workflows

All workflows

Docker Image CI

Python Package using Conda

## All workflows

Showing runs from all workflows

Filter workflow runs

50 workflow runs

Event ▾

Status ▾

Branch ▾

Actor ▾

✓ add test

Docker Image CI #16: Commit 89948a5 pushed by jayzhan211

master

8 minutes ago  
2m 2s

...

✓ add test

Python Package using Conda #34: Commit 89948a5 pushed by jayzhan211

master

8 minutes ago  
1m 28s

...

✓ add integration test

Python Package using Conda #33: Commit 985e745 pushed by jayzhan211

master

22 minutes ago  
1m 11s

...

✓ add integration test

Docker Image CI #15: Commit 985e745 pushed by jayzhan211

master

22 minutes ago  
2m 4s

...

✓ update company count

Python Package using Conda #32: Commit 0f2be32 pushed by jayzhan211

master

2 hours ago  
1m 14s

...


# Pytest Coverage

1. Run pytest-cov
2. Read pytest-coverage.txt

```
31     - name: Build coverage file
32       run: |
33         pytest --junitxml=pytest.xml --cov-report=term-missing:skip-covered --cov=./crawler tests/ | tee pytest-coverage.txt
34     - name: Pytest coverage comment
35       uses: MishaKav/pytest-coverage-comment@main
36       with:
37         pytest-coverage-path: ./pytest-coverage.txt
38         junitxml-path: ./pytest.xml
```

# Pytest Coverage





Github-actions Coverage  
Report  
TOTAL COVERAGE: 32%

 **github-actions** bot commented on [89948a5](#) 14 minutes ago

**Coverage** 32%

▼ Coverage Report

File	Stmts	Miss	Cover	Missing
<b>crawler</b>				
<a href="#">GoogleCrawler.py</a>	144	106	26%	<a href="#">20-37, 40-45, 49-52, 56-62, 66-75, 81-94, 103-111, 116-150, 153-171, 200-222</a>
<a href="#">GoogleCrawlerTest.py</a>	49	49	0%	<a href="#">1-76</a>
<a href="#">worker.py</a>	59	18	69%	<a href="#">20-27, 30-35, 40-41, 79-80</a>
<b>TOTAL</b>	<b>253</b>	<b>173</b>	<b>32%</b>	

Tests	Skipped	Failures	Errors	Time
7	0 	0 	0 	2.207s 



# Pytest Coverage

Github-actions Coverage Report

TOTAL COVERAGE: 32%

MISSING: Line Not Covered

```
16 rds = Redis('redis', 6379)
17 rq = Queue(connection=rds)
18
19 def work(url, timestamp):
20     response = get_source(url)
21     if response is None:
22         return None
23     soup = html_parser(response.text)
24     if soup is None:
25         return None
26     orignal_text = html_getText(soup)
27     return company_count(orignal_text, timestamp)
28
29 def get_source(url):
30     try:
```

GoogleCrawlerTest.py				
worker.py	59	18	69%	20-27, 30-35, 40-41, 79-80
TOTAL	253	173	32%	

# Github Action: Continuous Push Docker Image

1. Run docker-compose to build Docker image
2. Login to Docker Hub
  - a. Set Docker Hub Username and Token to github
  - b. Run login-action
3. Push Docker image

```
steps:  
- uses: actions/checkout@v3  
- name: Build the Docker image  
  run: docker-compose up -d  
  
- name: Login to DockerHub  
  uses: docker/login-action@v2  
  with:  
    username: ${ secrets.DOCKER_HUB_USERNAME }  
    password: ${ secrets.DOCKER_HUB_TOKEN }  
  
- name: Push Docker Image  
  run: docker-compose push
```

# Option

(高承翰)

# App Monitor: RQ dashboard

RQ dashboard

← → ↻ 🏠 ⚠️ 不安全 | 34.80.185.32:5000/rq/

🔍 ☆ ⚙️ 🗖️ 🚫

## RQ Instances

Select below the RQ instance that you want to observe.

redis://redis:6379 ▼

### Queues (toggle)

This list below contains all the registered queues with the number of jobs currently in the queue. Select a queue from above to view all jobs currently pending on the queue.

Queue	Jobs
📄 default	0
📄 failed	18

### Workers (toggle)

4 workers registered

State	Worker	Queues
II	162f79943e514f25865e164a7c498dc9	default
II	2158d31f9c154a9297b78b071911ef23	default
II	3a4dac1a6cbd40a3b0851957171610bf	default
II	eefaf5d81d9842d382f370d7a57f77fc	default

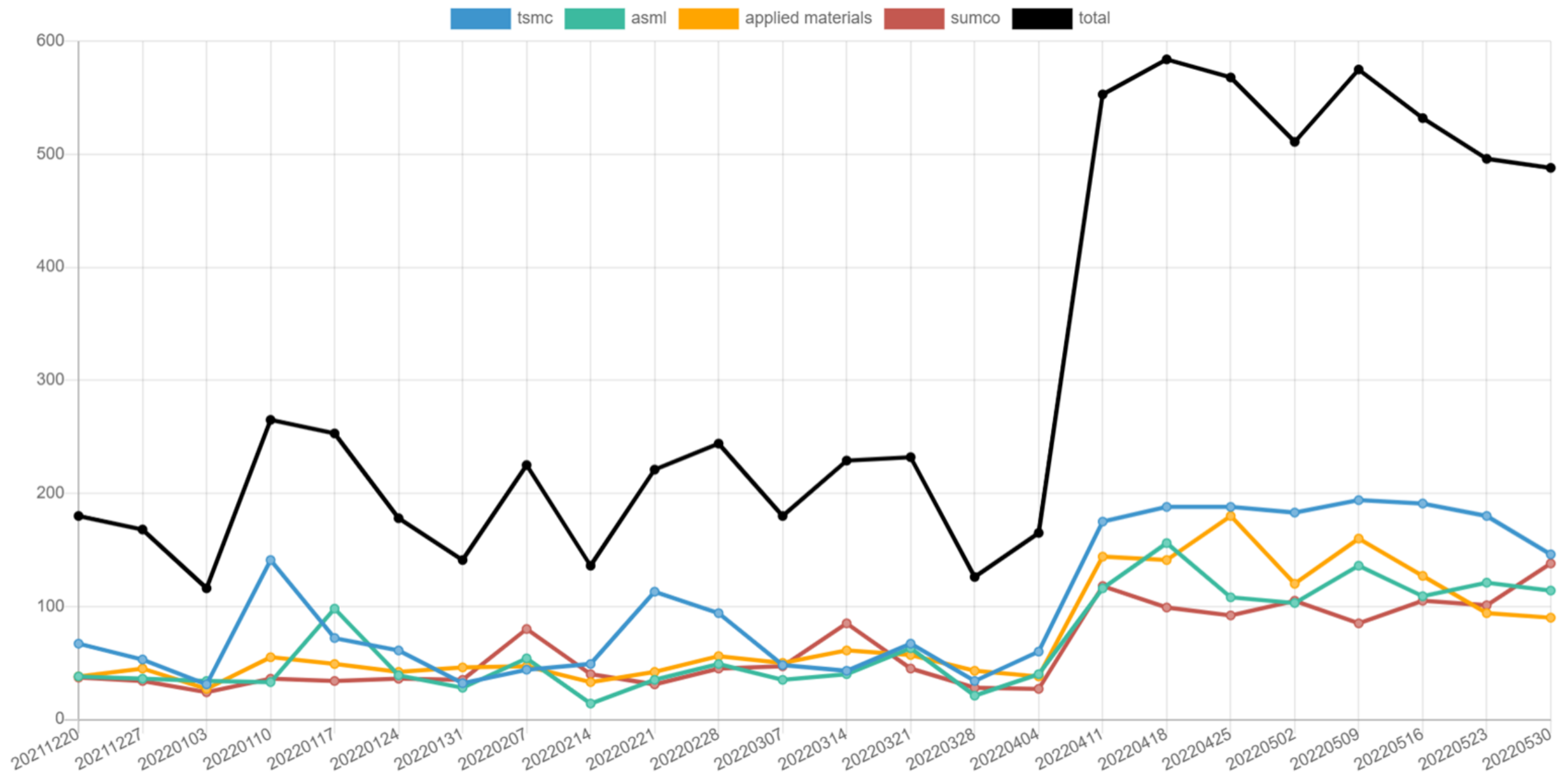
## Jobs on failed

This list below contains all the registered jobs on queue failed, sorted by age (oldest on top).

🔄 Requeue All

Name	Age	Actions
<div>📄 worker.work('https://m.toutiao.com/search?keyword=asm%E8%82%A1%E7%A5%A8&amp;pd=synthesis&amp;s..., '2022-05-23') from default</div> <div>aed0114c-c01b-4dc7-9848-1596b922224f</div> <div>Failed 2 hours ago</div> <div>Traceback (most recent call last): File "/usr/local/lib/python3.10/site-packages/rq/worker.py", line 1061, in perform_job rv = job.perform() File "/usr/local/lib/python3.10/site-packages/rq/job.py", line 821, in perform self._result = self._execute() File "/usr/local/lib/python3.10/site-packages/rq/job.py", line 844, in _execute result = self.func(*self.args, **self.kwargs) File "/tsmc-project/crawler/worker.py", line 19, in work response = get_source(url)</div>	2 hours ago	<div>🔄 Requeue</div> <div>✖ Cancel</div>

# App Monitor -- week url counts



# App Monitor: worker performance

```
Testing performance of 4 workers processing urls from 2022-05-30 to 2022-06-05
sleeping...
[Check][URL] URL : https://www.google.com/search?q=tsmc before:2022-06-05 after:2022-05-30&start=0&num=100&hl=en
sleeping...
[Check][URL] URL : https://www.google.com/search?q=tsmc before:2022-06-05 after:2022-05-30&start=100&num=100&hl=en
sleeping...
[Check][URL] URL : https://www.google.com/search?q=tsmc before:2022-06-05 after:2022-05-30&start=200&num=100&hl=en
Get 148 urls for tsmc from 2022-05-30 to 2022-06-05
sleeping...
[Check][URL] URL : https://www.google.com/search?q=asml before:2022-06-05 after:2022-05-30&start=0&num=100&hl=en
sleeping...
[Check][URL] URL : https://www.google.com/search?q=asml before:2022-06-05 after:2022-05-30&start=100&num=100&hl=en
sleeping...
[Check][URL] URL : https://www.google.com/search?q=asml before:2022-06-05 after:2022-05-30&start=200&num=100&hl=en
Get 110 urls for asml from 2022-05-30 to 2022-06-05
sleeping...
[Check][URL] URL : https://www.google.com/search?q=applied materials before:2022-06-05 after:2022-05-30&start=0&num=100&hl=en
sleeping...
[Check][URL] URL : https://www.google.com/search?q=applied materials before:2022-06-05 after:2022-05-30&start=100&num=100&hl=en
Get 92 urls for applied materials from 2022-05-30 to 2022-06-05
sleeping...
[Check][URL] URL : https://www.google.com/search?q=sumco before:2022-06-05 after:2022-05-30&start=0&num=100&hl=en
sleeping...
[Check][URL] URL : https://www.google.com/search?q=sumco before:2022-06-05 after:2022-05-30&start=100&num=100&hl=en
sleeping...
[Check][URL] URL : https://www.google.com/search?q=sumco before:2022-06-05 after:2022-05-30&start=200&num=100&hl=en
Get 137 urls for sumco from 2022-05-30 to 2022-06-05
487 jobs doned in 195.77 seconds, throughput:2.49 job/s
```

# Database Schema

tsmc\_project.word\_count

Documents	Aggregations	Schema	Explain Plan
FILTER { field: 'value' }			
ADD DATA VIEW			
<pre>{   "_id": ObjectId('6297227e0adf99ab37b7713b'),   "Date": "2022-05-09",   "Company": "tsmc",   "Word_Count": 648 }</pre>			
<pre>{   "_id": ObjectId('6297227e0adf99ab37b7713c'),   "Date": "2022-05-09",   "Company": "applied materials",   "Word_Count": 264 }</pre>			
<pre>{   "_id": ObjectId('6297227e0adf99ab37b7713d'),   "Date": "2022-05-09",   "Company": "sumco",   "Word_Count": 103 }</pre>			

tsmc\_project.url\_count

Documents	Aggregations	Schema	Explain Plan
FILTER { field: 'value' }			
ADD DATA VIEW			
<pre>{   "_id": ObjectId('629724b098d2b9f1794fb125'),   "Date": "2022-05-09",   "Company": "asml",   "Url_Count": 126 }</pre>			
<pre>{   "_id": ObjectId('629724b098d2b9f1794fb126'),   "Date": "2022-05-09",   "Company": "tsmc",   "Url_Count": 174 }</pre>			
<pre>{   "_id": ObjectId('629724b098d2b9f1794fb127'),   "Date": "2022-05-09",   "Company": "sumco",   "Url_Count": 128 }</pre>			

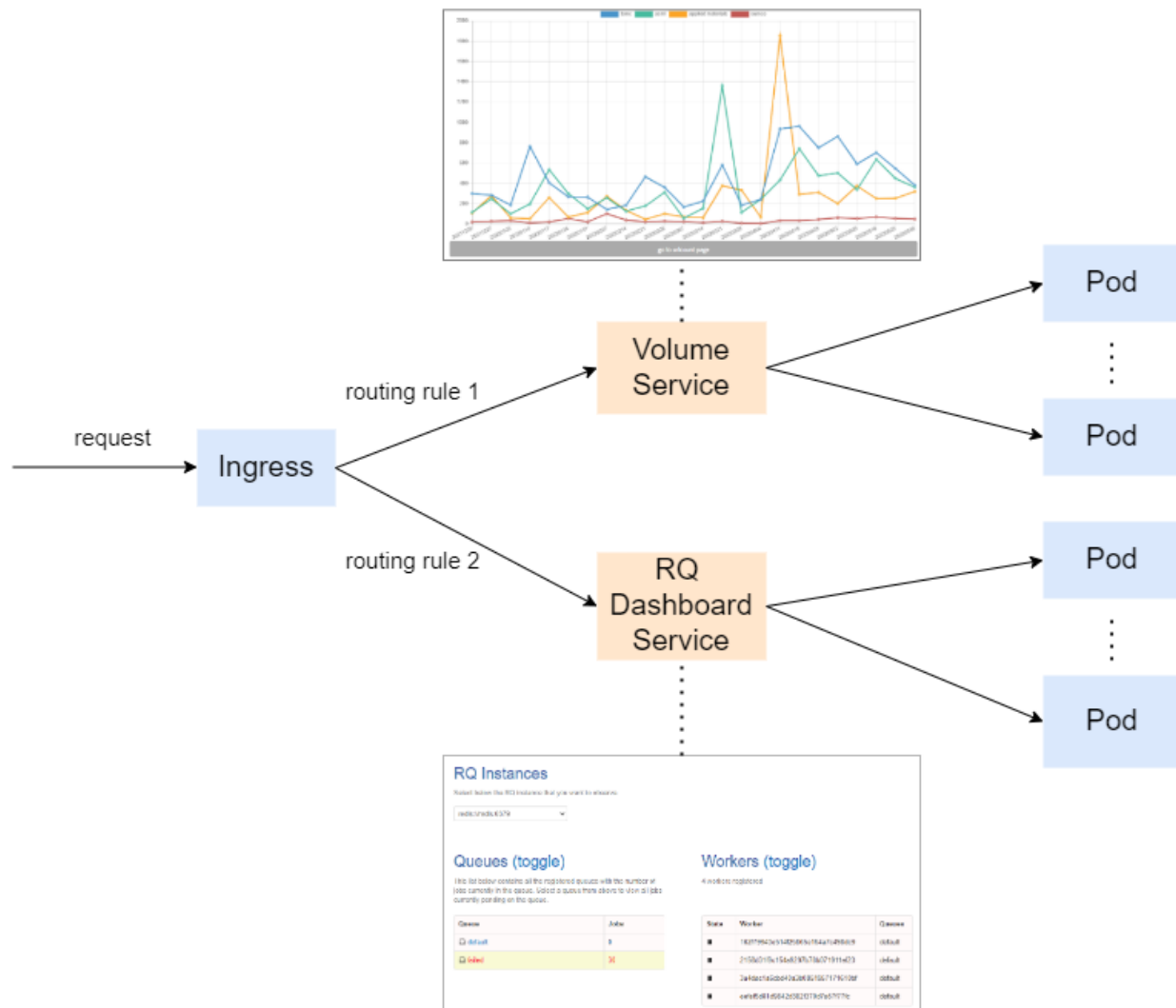


# Code Contribution

	Type	Sample Code	Our Application
1	Google Search	Search by default	Search by week
2	Word Count	✓	✓
3	URL Count	✗	✓
4	Task Queue	✗	✓
5	Testing	Unit test	Unit test + Integration test
6	Database	✗	MongoDB
7	Display	Excel	Web Page
8	Evaluate Performance	✗	✓
9	Update	✗	Every week



# Ingress



# Discussion

(廖家鴻)

# GKE Autopilot Deployment

Google Cloud Platform nycu-lab-10

Kubernetes 叢集 + 建立 + 部署 重新整理

總覽 成本最佳化

篩選 輸入屬性名稱或值

<input type="checkbox"/>	狀態	名稱 ↑	位置	模式	節點數量	vCPU 總數	記憶體總量	通知	標籤
<input type="checkbox"/>	✓	nycu-lab-10-cluster	asia-east1-b	Standard	0	0	0 GB	⚠ Pod 無法排程	—
<input type="checkbox"/>	✓	tsmc-cluster	asia-east1	Autopilot		4.5	16 GB		—

Pods

```
mg168allen@cloudshell:~ (nycu-lab-10) $ kubectl get pod -o wide
W0606 14:04:38.351914 536 gcp.go:120] WARNING: the gcp auth plugin is deprecated in v1.22+, unavailable in v1.25+; use gcloud instead.
To learn more, consult https://cloud.google.com/blog/products/containers-kubernetes/kubectl-auth-changes-in-gke
NAME                                READY   STATUS    RESTARTS   AGE   IP              NODE                                NOMINATED NODE   READINESS GATES
crawler-27574985-fdg74              0/1     Completed 0           6h59m 10.28.0.195     gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v <none>           <none>
frontend-58699f8dc4-8fwvr           1/1     Running   0           25h   10.28.0.236     gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v <none>           <none>
frontend-58699f8dc4-pbj92           1/1     Running   0           25h   10.28.0.237     gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v <none>           <none>
frontend-58699f8dc4-vz84j           1/1     Running   0           25h   10.28.0.238     gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v <none>           <none>
mongo-56cf6b45fc-tnpd6              1/1     Running   0           27h   10.28.1.198     gk3-tsmc-cluster-default-pool-13b292fa-vttf <none>           <none>
nodeserver-f7458bd5b-r27vf          1/1     Running   0           27h   10.28.2.68      gk3-tsmc-cluster-default-pool-c11a1389-fn59 <none>           1/1
redis-9b88c854c-mdkrq               1/1     Running   0           27h   10.28.0.93      gk3-tsmc-cluster-default-pool-c11a1389-7qfk <none>           <none>
worker-7f47c8d444-4h5qd             1/1     Running   0           27h   10.28.0.208     gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v <none>           <none>
worker-7f47c8d444-7vgr6             1/1     Running   0           27h   10.28.0.207     gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v <none>           <none>
worker-7f47c8d444-gdhvj             1/1     Running   0           27h   10.28.0.210     gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v <none>           <none>
worker-7f47c8d444-x2bc6             1/1     Running   0           27h   10.28.0.209     gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v <none>           <none>
```

Nodes

```
mg168allen@cloudshell:~ (nycu-lab-10) $ kubectl get node
W0606 14:18:12.326785 568 gcp.go:120] WARNING: the gcp auth plugin is deprecated in v1.22+, unavailable in v1.25+; use gcloud instead.
To learn more, consult https://cloud.google.com/blog/products/containers-kubernetes/kubectl-auth-changes-in-gke
NAME                                STATUS   ROLES    AGE   VERSION
gk3-tsmc-cluster-default-pool-13b292fa-fw6b Ready    <none>   2d23h v1.22.8-gke.201
gk3-tsmc-cluster-default-pool-13b292fa-vttf Ready    <none>   28h    v1.22.8-gke.201
gk3-tsmc-cluster-default-pool-c11a1389-7qfk Ready    <none>   2d23h v1.22.8-gke.201
gk3-tsmc-cluster-default-pool-c11a1389-fn59 Ready    <none>   28h    v1.22.8-gke.201
gk3-tsmc-cluster-nap-19qcyna4-744da966-4k7v Ready    <none>   28h    v1.22.8-gke.201
```

# About Testing

## □ The different types of tests

- Unit Test (單元測試)
- Integration Test (整合測試)
- End to End Testing (E2E測試)
- Acceptance testing (驗收測試)
- Performance testing

# About Monitor

## Monitoring Methodologies

- There are three common lists or methodologies.
  - [USE Method](#) (for resource): **Utilization, Saturation, and Errors**
  - [RED Method](#) (for service): **Rate, Errors, and Duration**
  - Google's four golden signals (from the [Google SRE book](#)): **Latency, Traffic, Errors, and Saturation**
- **Latency** — Response time, including queue/wait time, in milliseconds.
- **Traffic (Rate)** — Request rate, in requests/sec
- **Errors** — Error rate, in errors/sec
  1. explicit failure: like HTTP 500 error
  2. implicit failure: like wrong or invalid content being returned
  3. policy-based failure: ex, over 10s should be considered error
- **Saturation** — How overloaded something is, which is related to utilization but more directly measured by things like queue depth (or sometimes concurrency). As a queue measurement, this becomes non-zero when you are saturated, often not much before.
- **Utilization** — How busy the resource or system is. Usually expressed 0–100% and most useful for predictions (as Saturation is probably more useful

# About Monitor

RQ dashboard link: <http://34.120.17.241/rq/>

## RQ Instances

Select below the RQ instance that you want to observe.

redis://redis:6379

▼

## Queues (toggle)

This list below contains all the registered queues with the number of jobs currently in the queue. Select a queue from above to view all jobs currently pending on the queue.

Queue	Jobs
 default	0
 failed	30

## Workers (toggle)

4 workers registered

State	Worker	Queues
II	162f79943e514f25865e164a7c498dc9	default
II	2158d31f9c154a9297b78b071911ef23	default
II	3a4dac1a6cbd40a3b0851957171610bf	default
II	eefaf5d81d9842d382f370d7a57f77fc	default

# Issues about CD-1

Google Cloud Platform

nycu-lab-10

搜尋 產品、資源、文件 (/)

服務帳戶

+ 建立服務帳戶

刪除

+ 管理存取權

重新整理

「nycu-lab-10」專案的服務帳戶

服務帳戶代表的是 Google Cloud 服務身分，例如在 Compute Engine VM 上所執行的程式碼、App Engine 應用程式，或在 Google 以外執行的系統，[進一步瞭解服務帳戶](#)。

可使用機構政策來保護服務帳戶，並封鎖較有風險的服務帳戶功能，例如自動授予身分與存取權管理權限、建立/上傳金鑰，或是完全封鎖建立服務帳戶的功能，[進一步瞭解服務帳戶機構政策](#)。

篩選 輸入屬性名稱或值

<input type="checkbox"/>	電子郵件	狀態	名稱 ↑	說明	金鑰 ID	金鑰建立日期	OAuth 2 用戶端 ID ?	動作
<input type="checkbox"/>	745228716152-compute@developer.gserviceaccount.com	✓	Compute Engine default service account		沒有任何金鑰		111734006040384010601 	⋮
<input type="checkbox"/>	gke-sa@nycu-lab-10.iam.gserviceaccount.com	✓	gke-sa		沒有任何金鑰		114121234872233455583 	⋮



# Issues about CD-2 (gke-cluster pods)

Google Cloud Platform

tsmc-proj

搜尋 產品、資源、文件 (/)

工作負載

重新整理

部署

刪除

叢集

命名空間

重設

SAVE

篩選


屬於系統物件 : False

篩選工作負載


<input type="checkbox"/>	名稱 ↑	狀態	類型	Pod	命名空間	叢集
<input type="checkbox"/>	crawler	✓ OK	Cron Job	0/0	default	autopilot-cluster-1
<input type="checkbox"/>	frontend	✓ OK	Deployment	3/3	default	autopilot-cluster-1
<input type="checkbox"/>	mongo	✓ OK	Deployment	1/1	default	autopilot-cluster-1
<input type="checkbox"/>	nodeserver	✓ OK	Deployment	1/1	default	autopilot-cluster-1
<input type="checkbox"/>	redis	✓ OK	Deployment	1/1	default	autopilot-cluster-1
<input type="checkbox"/>	worker	! 3 of 4 updated replicas available - Unschedulable and 1 more issue	Deployment	3/4	default	autopilot-cluster-1





# Issues about CD-3 (github-action workflow)


 AaronKFC / tsmc-project Public

forked from CHKao777/tsmc-project


 Pin


 Watch 0


 Fork 3


 Star 0


<> Code


 Pull requests


 Actions

 Projects

 Wiki

 Security


 Insights


 Settings


Workflows

New workflow

All workflows

 Build and Deploy to GKE










 Docker Image CI

 Python Package using Conda

## All workflows

Showing runs from all workflows

Filter workflow runs

27 workflow runs	Event	Status	Branch	Actor
<div> Update GKE-CD.yml</div> <div>Docker Image CI #12: Commit 0082736 pushed by AaronKFC</div> <div>master</div> <div> 16 minutes ago</div> <div> 2m 12s</div> <div>...</div>				
<div> Update GKE-CD.yml</div> <div>Python Package using Conda #12: Commit 0082736 pushed by AaronKFC</div> <div>master</div> <div> 16 minutes ago</div> <div> 1m 17s</div> <div>...</div>				
<div> Update GKE-CD.yml</div> <div>Build and Deploy to GKE #3: Commit 0082736 pushed by AaronKFC</div> <div>master</div> <div> 16 minutes ago</div> <div> 48s</div> <div>...</div>				

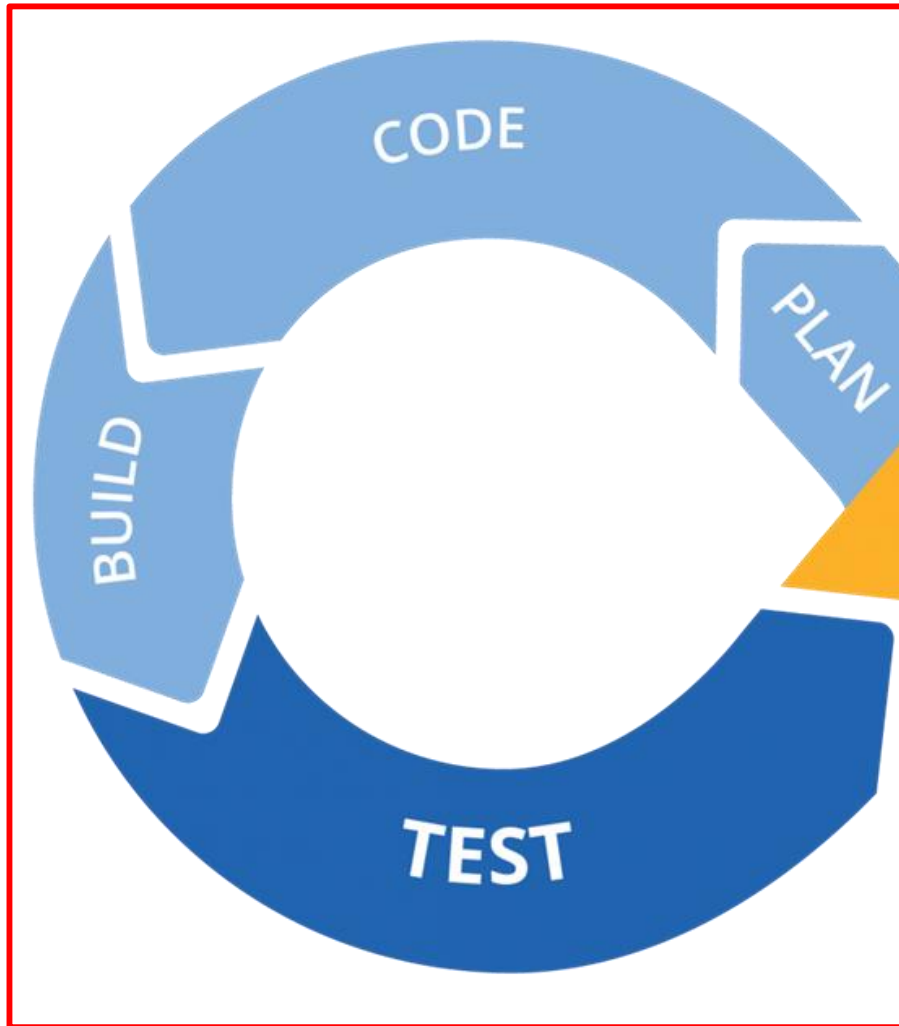
Service link: <https://github.com/AaronKFC/tsmc-project/actions>

# Conclusion

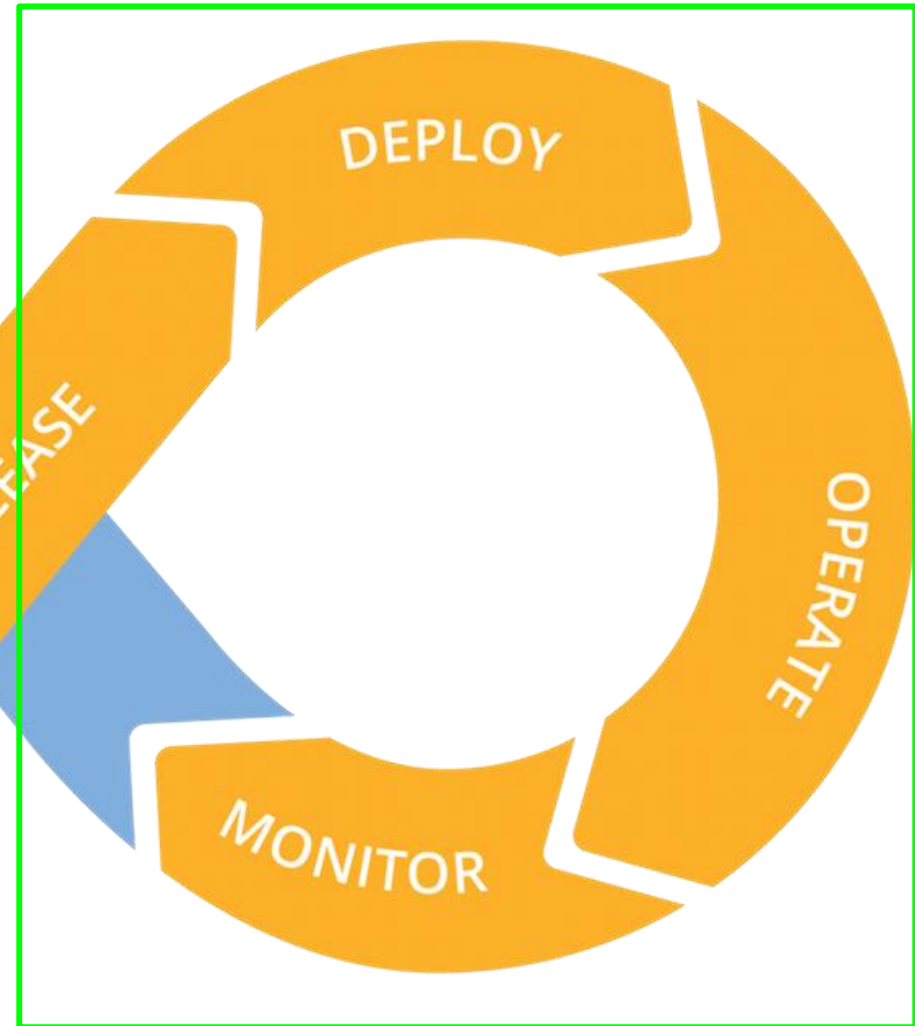
# Conclusion

Type	Category	Remark
Basic	Upload codes to Github	version control
	Deploy to GKE	use GKE Autopilot to build cluster
	Digital Business: Internet volume changes	word count weekly from Dec. 2021 to May 2022
	Application Architecture	GoogleCrawler, Workers, Web Server...
	Code Optimization	architecture optim, improve reliability, SRP
	Testing Coverage	unit test, integration test
	CI	Github Action, Pytest coverage
	CD	1.Github Action push Docker Image to Dockerhub 2.we don't have permission of GKE service account
Option	Application Monitor	RQ dashboard, url count, worker performance
	Use Database	schema, MongoDB, Redis
	Code Contribution	relative to sample codes
	Ingress	<a href="http://34.120.17.241/">http://34.120.17.241/</a>

# DevOps Sharing



CI



CD

**Thank You**

Q & A



# Final Project Evaluation (1/2)

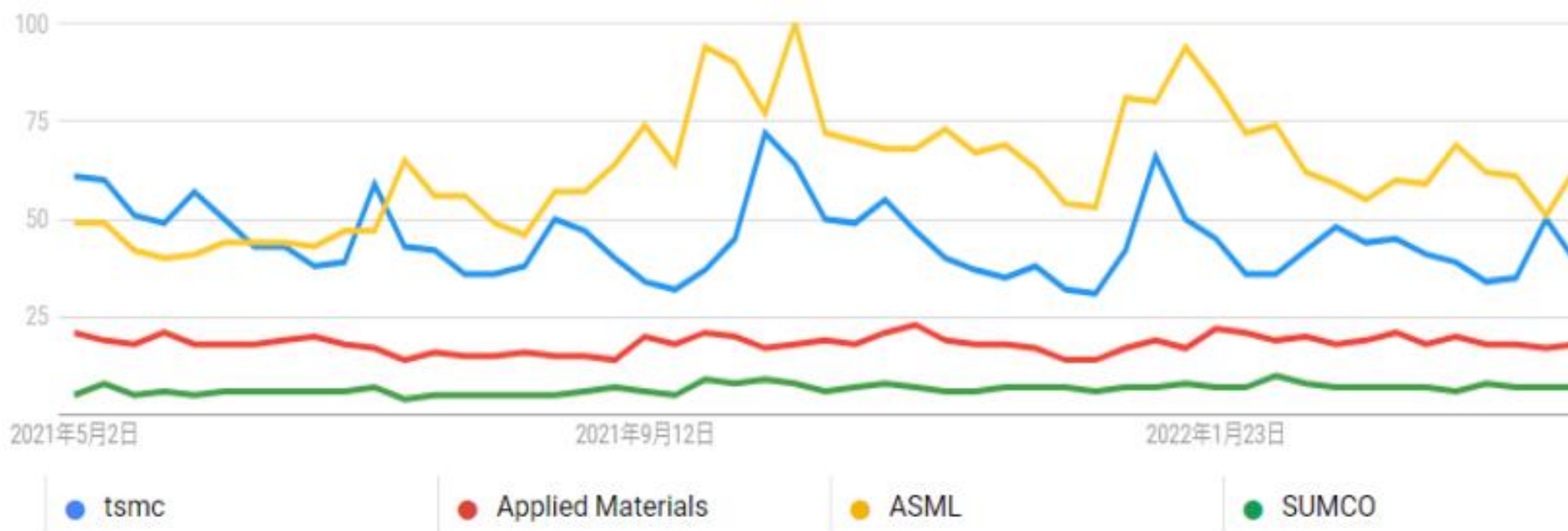
Type	Category	Score %	Deliverables & Evaluation	Remark
Basic (100%)	Application meet requirements	60%	Upload codes to Github and deploy to GKE	Provide (4/30) application codes with unit test, deployment docker image
	Testing Coverage (Junit, Pytest .. With Cobertura)			
	Application Architecture	20%	Provide PPT wi/ final presentation	Provide sample of architecture
	Code Optimization (adopting design patterns to keep code maintainable, extensible or improve reliability or faster trouble shooting) , CI/CD	20%	1) Upload codes to Github and deploy to GKE 2) Provide PPT wi/ final presentation	Enhance and optimize template codes
Option (30%)	Application Monitor (Metrics, Approach ..)	10%	Provide PPT wi/ final presentation	5/11 lectures
	Code Quality (configurable design, password encryption..) + 增加 unit test/logging, Use DB+ code contribution	20%	Provide PPT wi/ final presentation	



# Introduction

## Digital Business

- 請透過網路爬蟲，找出台積電與3家供應商的網路聲量變化。
  - Applied Materials (應用材料)
  - ASML (艾司摩爾)
  - SUMCO (勝高株式會社)

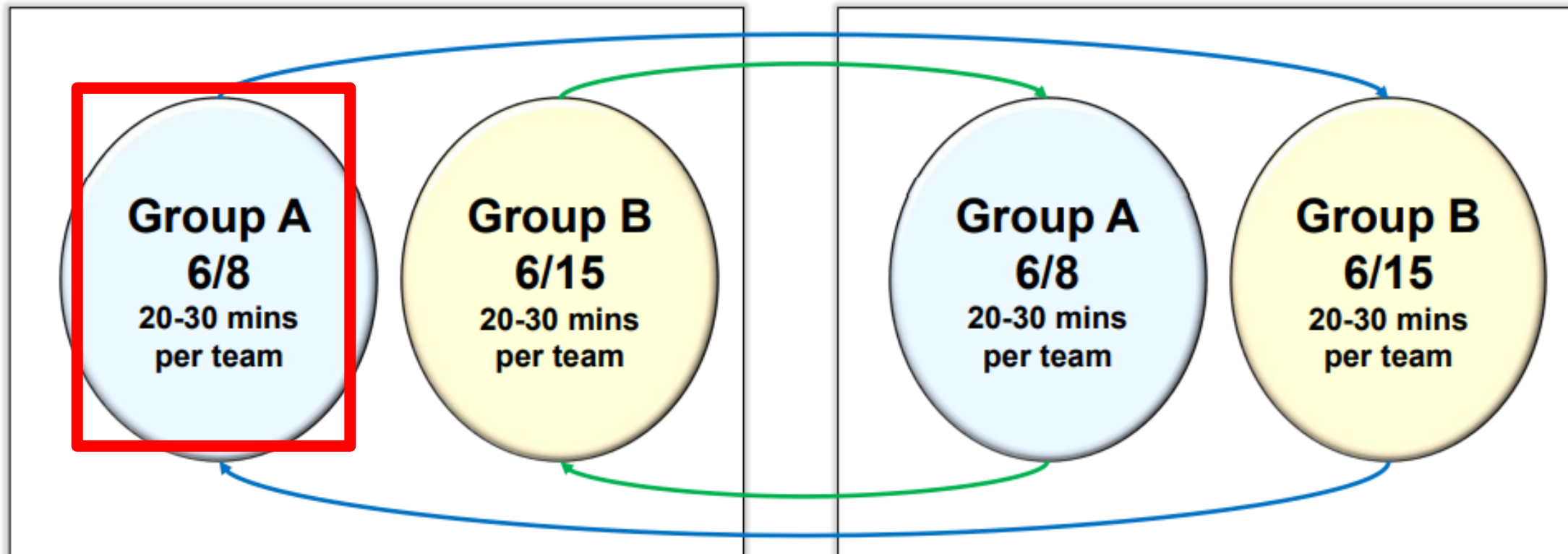


## Final Project Evaluation (2/2)

- (Option) Final Project 準備 + 助教諮詢 at 6/1
- Github code frozen at 6/7
- Two presentations at 6/8, 6/15 and welcome to join other group at free week.

### Digital Business

### Smart Manufacture



# Digital Business 報告組別及順序

- Sample code on GitHub: [https://github.com/NLPting/TSMC\\_NYCU\\_SAMPLE](https://github.com/NLPting/TSMC_NYCU_SAMPLE)
- 如要交換報告順序，請在 6/7 前找好要交換的組別並通知助教

## 6/8

組別	組長
4	簡言安
10	遊智鈞
13	王志嘉
16	劉祐誠
18	蘇冠安
20	陳毅

## 6/15

組別	組長
11	林奕宏
12	張嫻嫻
14	陳思羽
15	林熙哲
21	林思齊
23	許志瑄

6/1 Final Project 會議連結 – [點此進入會議](#)  
Digital Business問題提問 – [連結](#)