

Portfolio – Allen Wu

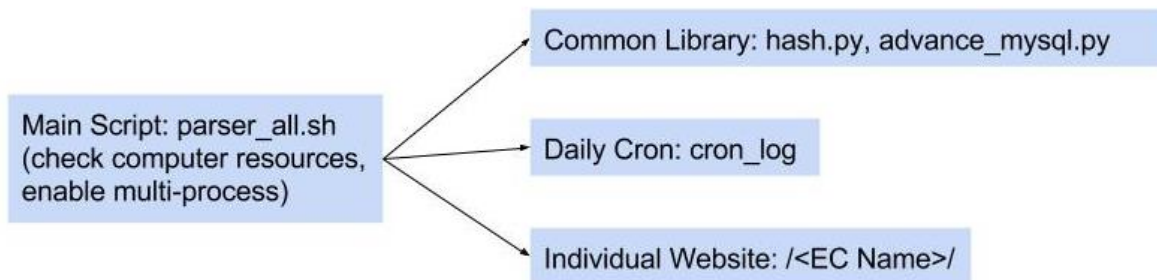
+886911677919•allencat850502@gmail.com•<https://github.com/allenwhale>

INTERNSHIP PROJECT

E-commerce web crawler

In order to collect E-commerce data for later analysis, our team wrote some crawlers to gather data from more than 10 different E-commerce websites. What's more, we also wrote a script to maintain these crawlers and ran them in multi-process with resources controlling.

- Product during internship of Industrial Technology Research Institute
- Techniques: Python, Shell Script, MySQL



PTT Crawler

In order to gather PTT articles for later trending analysis, I wrote a package in Python to retrieve data from PTT website. By the way, to develop more convenient, it can be install as a package of Python and get articles iteratively.

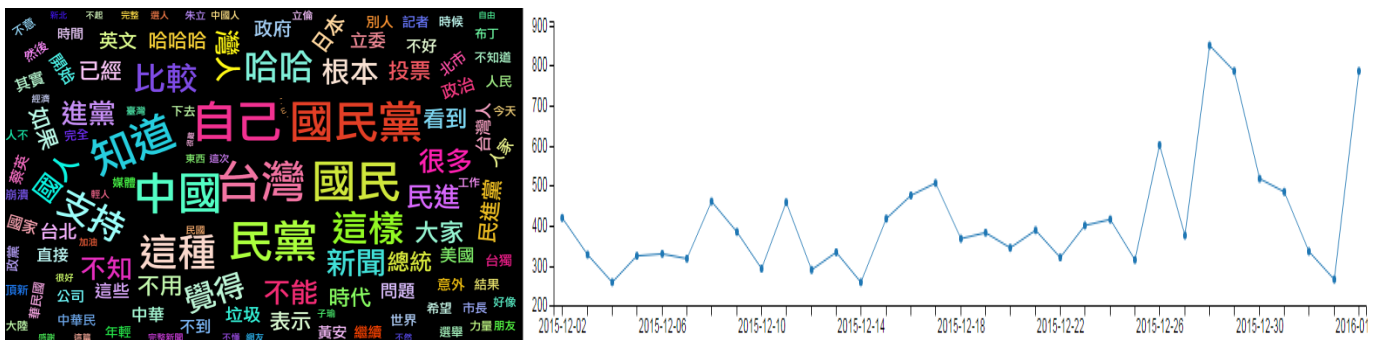
- Product during internship of Industrial Technology Research Institute
- Techniques: Python, MySQL
- Github Repo: https://github.com/allenwhale/ptt_crawler



Ngrams

To find out the popular words of articles or news, we calculate the ngrams of words. In order to improve the performance, we programmed it in Apache Spark which can run in multi-process. By the way, we show the trending result in web with c3.js.

- Techniques: tornado(Python), Apache Spark, Shell Script



NCTUOJ PROJECT

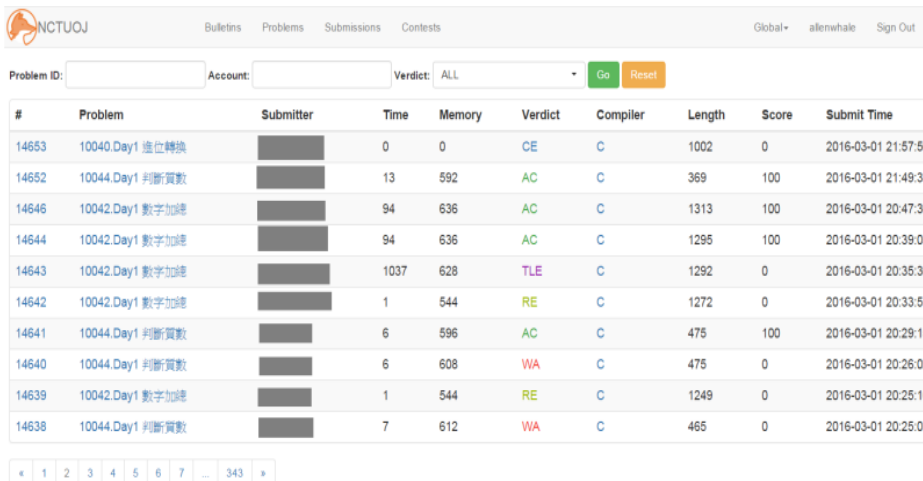
I and my friend started this project for improving our skills in beginning. After we developed a beta version, version contacted with teachers and put this project in real course of our school.

This project is divided into four parts, including web, judge center, judge client, Dockerfile.

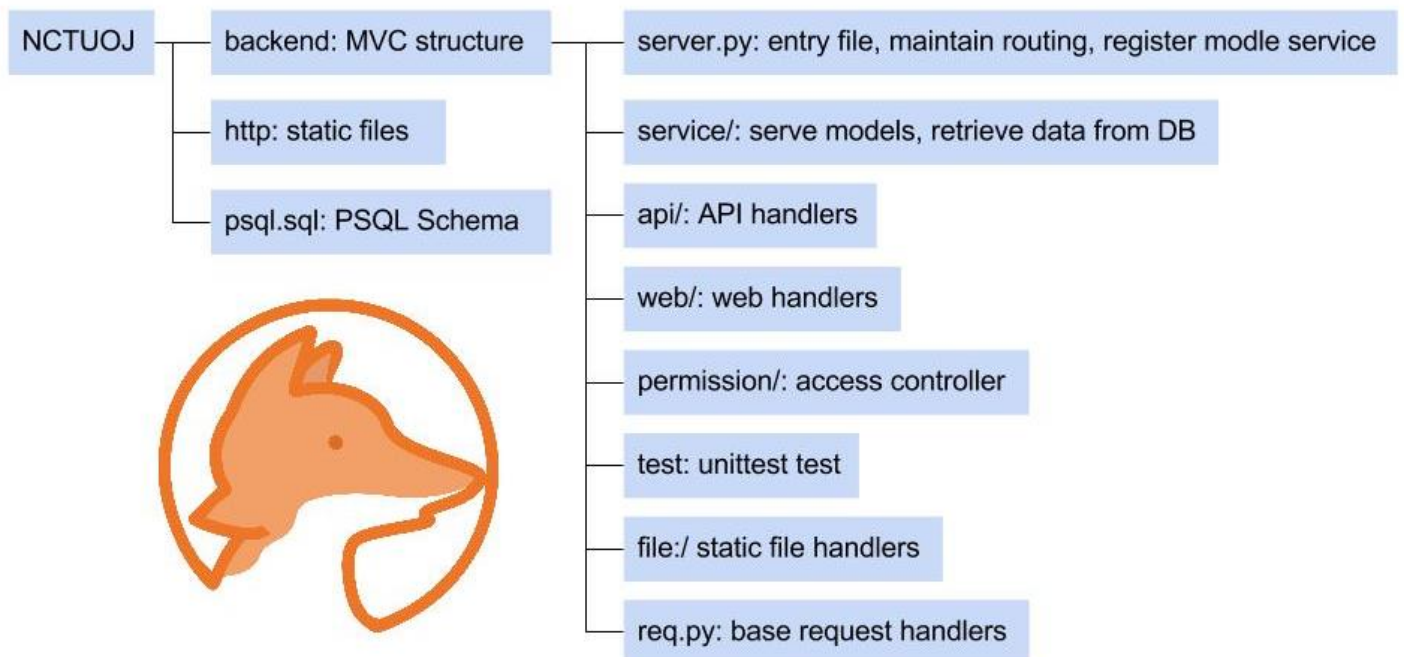
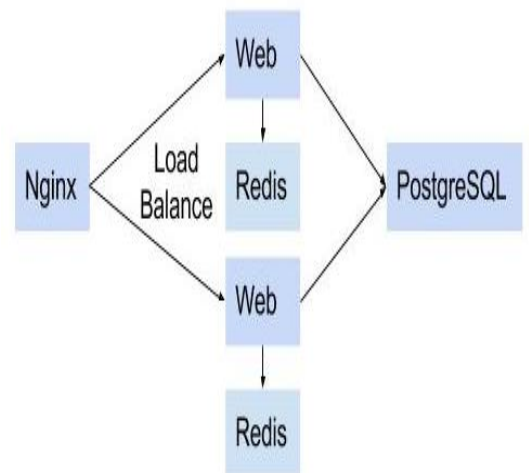
Web

Website developed in tornado, a Python web framework and asynchronous networking library. We created our own MVC structure, for developed more gracefully. We selected PostgreSQL as database system, and Redis as local cache to reduce SQL query time. In addition, we used Nginx for load balance to deal with more requests.

- Techniques: Nginx, Python, PostgreSQL, Redis
- Github Repo: <https://github.com/tocknicsu/nctuoj>



#	Problem	Submitter	Time	Memory	Verdict	Compiler	Length	Score	Submit Time
14653	10040.Day1 進位轉換		0	0	CE	C	1002	0	2016-03-01 21:57:56
14652	10044.Day1 判斷質數		13	592	AC	C	369	100	2016-03-01 21:49:33
14646	10042.Day1 數字加總		94	636	AC	C	1313	100	2016-03-01 20:47:30
14644	10042.Day1 數字加總		94	636	AC	C	1295	100	2016-03-01 20:39:05
14643	10042.Day1 數字加總		1037	628	TLE	C	1292	0	2016-03-01 20:35:30
14642	10042.Day1 數字加總		1	544	RE	C	1272	0	2016-03-01 20:33:51
14641	10044.Day1 判斷質數		6	596	AC	C	475	100	2016-03-01 20:29:15
14640	10044.Day1 判斷質數		6	608	WA	C	475	0	2016-03-01 20:26:02
14639	10042.Day1 數字加總		1	544	RE	C	1249	0	2016-03-01 20:25:10
14638	10044.Day1 判斷質數		7	612	WA	C	465	0	2016-03-01 20:25:01



Judge Center

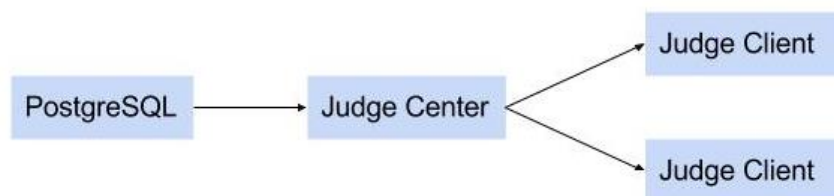
We maintain a table for those waiting for judging. After the center started, it will retrieve data from the table and dispatch tasks to clients

- Techniques: Python
- Github Repo: <https://github.com/Tocknicsu/judge-center/>

Judge Client

Judge client can be started and connect to judge center easily. Admin can increase or decrease the number of client according to the loading of requests. We selected CMS-isolate as a sandbox to judge the test file from users. The sandbox can prevent the damages from the users' programs simply.

- Techniques: Python, C/C++
- Github Repo: <https://github.com/Tocknicsu/judge-client>

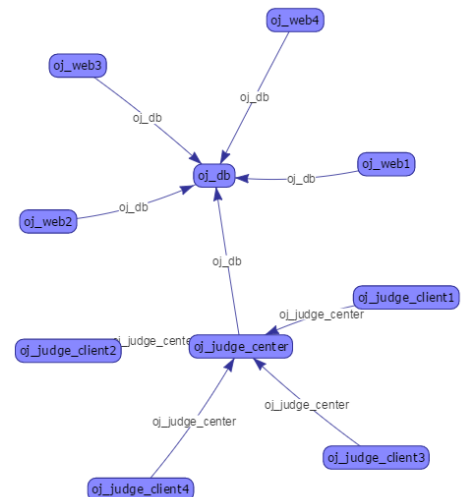


Dockerfile

In order to deploy NCTUOJ easily, we wrote Dockerfile, so we can create docker containers in only several commands.

- Techniques: Dockerfile, Shell Script
- Github Repo: https://github.com/allenwhale/nctuoj_docker

Action	Name	Image	Command	Created	Status
<input type="checkbox"/>	oj_judge_client4	judge-client	bash ./judge.client.sh	Mon Feb 29 2016	Up 46 hours
<input type="checkbox"/>	oj_judge_client3	judge-client	bash ./judge.client.sh	Mon Feb 29 2016	Up 46 hours
<input type="checkbox"/>	oj_judge_client1	judge-client	bash ./judge.client.sh	Mon Feb 29 2016	Up 2 days
<input type="checkbox"/>	oj_judge_client2	judge-client	bash ./judge.client.sh	Mon Feb 29 2016	Up 2 days
<input type="checkbox"/>	oj_judge_center	judge-center	bash ./judge.center.sh	Mon Feb 29 2016	Up 2 days
<input type="checkbox"/>	oj_web4	web	bash ./web.sh	Sun Feb 28 2016	Up 2 days
<input type="checkbox"/>	oj_web3	web	bash ./web.sh	Sun Feb 28 2016	Up 2 days
<input type="checkbox"/>	oj_web2	web	bash ./web.sh	Sun Feb 28 2016	Up 2 days
<input type="checkbox"/>	oj_web1	web	bash ./web.sh	Sun Feb 28 2016	Up 2 days
<input type="checkbox"/>	oj_db	db	/docker-entrypoint.sh postgres	Sun Feb 28 2016	Up 2 days



Planned Feature

Website

- Use pgpool to enhance PostgreSQL performance.
- Make the UI/UX more user-friendly.

Judge Center

- Automatically increase or decrease number of judge client.

Judge Client

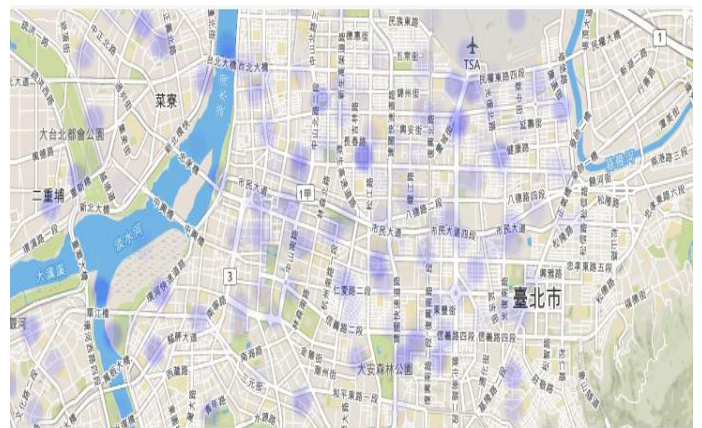
- Improve to accuracy of execution time and memory usage of test program.

COMPETITIONS

2016 Mei-Chu Hackathon

in this competition, we created a new way of third party payment through QR code. We used E.Sun Bank's API to complete online transactions, and enabled people to create and micropayment faster. Because our method only need to scan the QR code, sellers or buyers DO NOT need to install any other application, this further increases the chance that it can be widespread. At last, we got the first place of E.Sun Banks group, the second place of whole competition and the UI/UX prize.

- Records
 - 2016 Mei-Chu Hackathon 2nd Place
 - E.Sun Bank group 1st place
 - UI/UX Prize
- Techniques: tornado(Python), HTML, CSS, JavaScript, MySQL
- Github Repo
 - https://github.com/kevchentw/nctu_hackathon
 - <https://github.com/kevchentw/VizMap>



ACM-ICPC

I and my friends attend ACM-ICPC as a group. We practice problem solving to improve our programming skill and our sight.

- Records
 - 2014 ACM-ICPC Asia Taichung Regional Contest 17th
 - 2015 ACM-ICPC Asia Taipei Regional Contest Rank 12 Third Prize
 - 2015 ACM-ICPC Asia Jakarta Regional Contest Rank 9
- Technique: C/C++, Java
- Github Repo
 - Codebook: <https://github.com/allenwhale/acm>
 - Past code: <https://github.com/allenwhale/code>