API Extension Data Exploration

Cindy Wang

INTRODUCTION

This report will evaluate the importance of 38 APIs and identify the most important application developer from 4 developers with the assumption below:

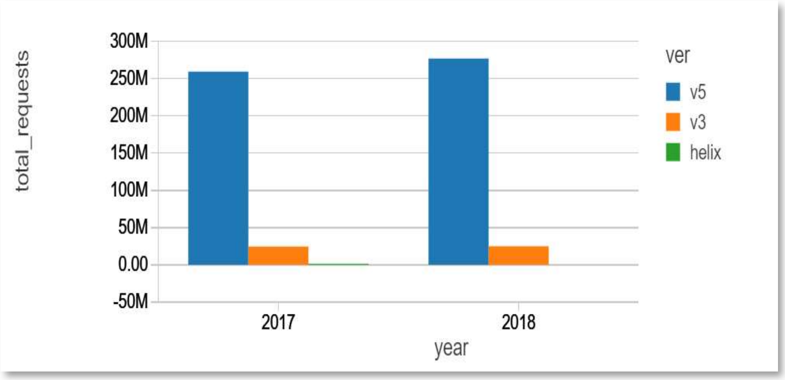
Assumption

Without a specific instruction for the definition of “importance”, I approach it from two aspects:

- The request frequencies
- The response speed

While the most frequently requested APIs are the most popular APIs in the ecosystem, the quickest responded APIs lay the foundation and process the simplest operations for the platform, like getting the ID of users, videos and channels. If exist, the best APIs should be both requested most frequently and processed most quickly. However, the number of such APIs will be too little to build a big picture. Therefore, for long-term planning, the most frequently requested APIs will be given priority in order to satisfy the needs of users; Meanwhile, we still cannot deny the importance of those basic APIs in the current stage. Therefore, this report will analyze from those two metrics at the same time and give priority to the frequencies of requests.

In addition, based on the context given, API has experienced three versions: V3, V5 and Helix (the most recent version). However, V3 and its third-party support have already been officially deprecated and though V5 became the default version in early 2019, it will also be removed in the short future. This report will consider the short-term as both V5 and Helix exist, the long-term as only Helix remains active. The chart below shows the use of different versions of APIs and we can see that the use of V3 API is little compared to V5. Thus, the assumption of not considering V3 is reasonable.



APPROACH AND FINDINGS

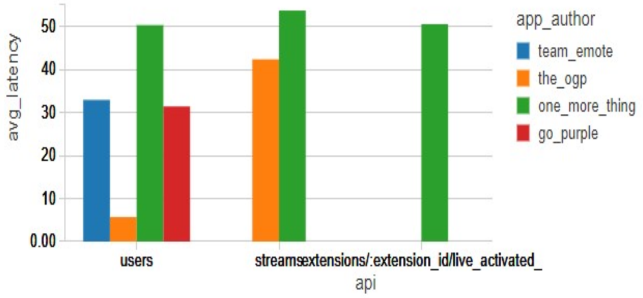
1. Short-term Planning (when both V5 and Helix exist)

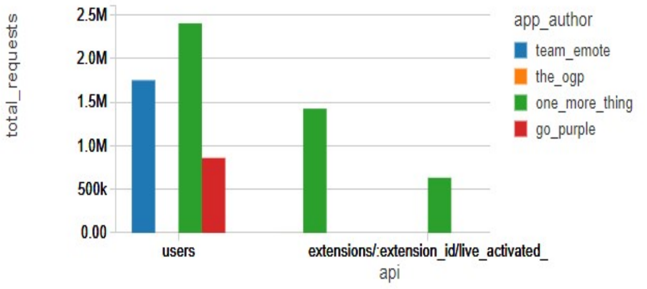
a. Both Most Frequently Requested and Quickly Responded APIs

The most important APIs are the ones that are frequently requested and can respond quickly at the same time, which are **streams(Helix), extensions/:extensionid/live_activated(v5) and users (v5).** The detailed information of each API is listed below:

ver	api	app_author	total_requests	avg_latency
helix	streams	one_more_thing	1426255	53.71
helix	streams	the_ogp	904	42.38
v5	extensions/:extension_id/live_activated_	one_more_thing	633833	50.55
v5	users	team_emote	1752226	32.94
v5	users	one_more_thing	2405148	50.33
v5	users	the_ogp	1142	5.71
v5	users	go_purple	860957	31.42

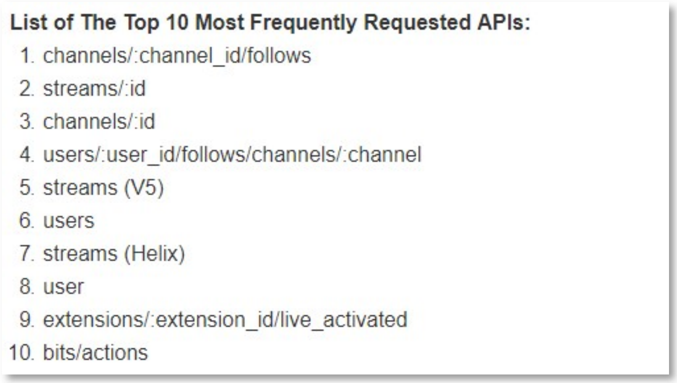
As we can see from the two charts below, the developers who design the quickest responded APIs are **the_ogp** and **go_purple** and the ones who design the most active APIs are **one_more_thing** and **team_emote.** That being said, different developers have different expertise and depending on the specific needs, we can choose the specific partners if needed. This report will give more evidence regarding this point.





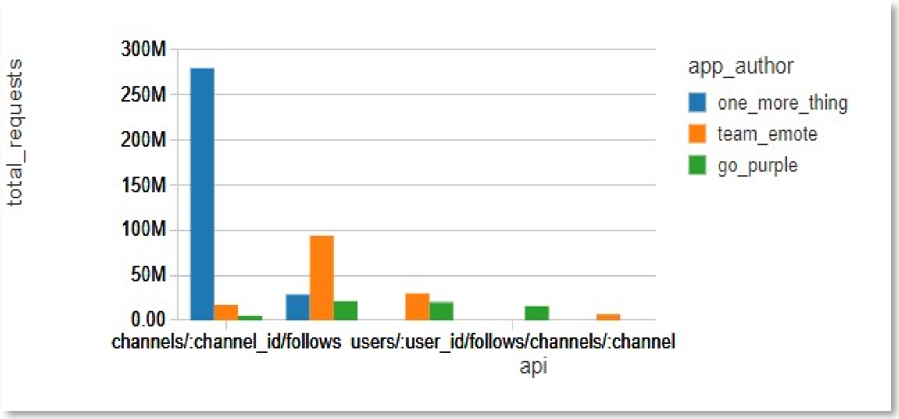
b. Most Frequently Requested APIs

In order to give a more comprehensive list of the most active APIs, I identify the Top 10 requested API as below:



List of The Top 10 Most Frequently Requested APIs:

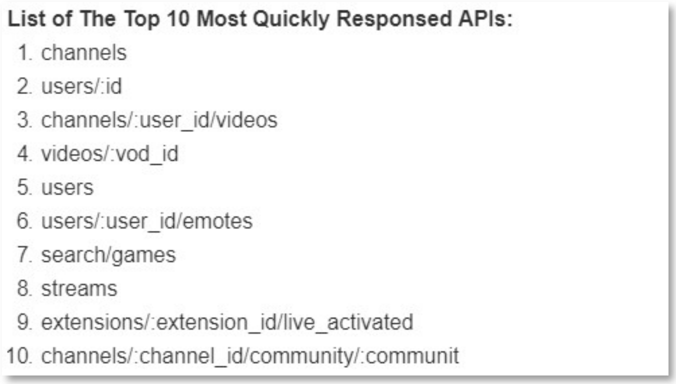
1. channels/:channel_id/follows
2. streams/:id
3. channels/:id
4. users/:user_id/follows/channels/:channel
5. streams (V5)
6. users
7. streams (Helix)
8. user
9. extensions/:extension_id/live_activated
10. bits/actions



In terms of the distribution of the developers of those APIs, the pattern is similar to what described above: the **one_more_thing** and **team_emote** have developed the most active APIs so far. Based on the chart, **one_more_thing** dominates the R&D of the most active APIs.

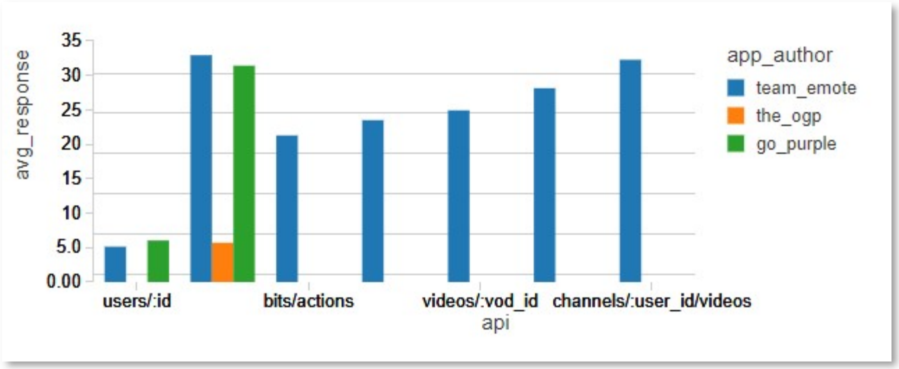
c. Quickest Responded APIs

Using the same approach, I summarize the quickest responded API and its predominant developers as show below. The result is consistent and **the_ogp** and **go_purple** turn out to be the developers of the quickest responded APIs.



List of The Top 10 Most Quickly Responsed APIs:

1. channels
2. users/:id
3. channels/:user_id/videos
4. videos/:vod_id
5. users
6. users/:user_id/emotes
7. search/games
8. streams
9. extensions/:extension_id/live_activated
10. channels/:channel_id/community/:communit



2. Long-term Exploration (Only Helix Remain Active)

In the long-term, as Helix APIs are still in the R&D stage, not many APIs are currently active on the platform. Among the developers of the only two APIs, the **one_more_thing** dominates the most active APIs and ranks the third in aspects of response speed. Without surprise, **the_ogp** is responsible for launching the quickest responded APIs. And **go_purple** has not launched any API in the newest version.

CONCLUSION

- **Which API endpoints are more important to continue to maintain?**

The top tier APIs are those which have the highest requests and respond the most quickly. That being said, they are the most fundamental APIs for the platform. The other important APIs all have different emphasis, either most requested or give the quickest responses. Depending on the specific needs, we can choose to maintain the specific APIs. However, it is important to notice that while some can satisfy the users' needs, some are the basic APIs needed to keep website's normal functions. Therefore, we should be careful in removing APIs.

Top tier APIs:

- users
- streams
- extensions/:extension_id/live_activated_

The second-tier important APIs that have the highest requests:

- channels/:channel_id/follows
- streams/:id
- channels/:id
- users/:user_id/follows/channels/:channel
- bits/actions

The third-tier important APIs that have the quickest responses:

- channels
- users/:id
- channels/:user_id/videos

- videos/:vod_id
- users/:user_id/emotes
- search/games
- channels/:channel_id/community/:communit

- **Which application developers are important to the developer ecosystem?**

From the table below, we can see that **one_more_thing** can produce the most popular APIs and **the_ogp** leads the development of fastest APIs. **Team_emote** and **go_purple** are balanced developers. However, **go_purple** does not involve in launching the most recent version of API. Therefore, in consideration of long-term partnership, product manager can choose from the remaining three developers.

Among the remaining three, **team_emote** will be a safe choice. But **one_more_thing** is the most far-sighted developer, who can quickly adapt to the changes. The total requests of its Helix APIs have the overwhelming advantage over others. If the web traffic allows, we could also consider the long-term partnership with **one_more_thing** and work together to improve its API performance.

	Avg Response Ranking	Total Requests Ranking	Involved in Helix API?
the_ogp	1	4	Yes
go_purple	2	3	No
team_emote	3	2	Yes
one_more_thing	4	1	Yes

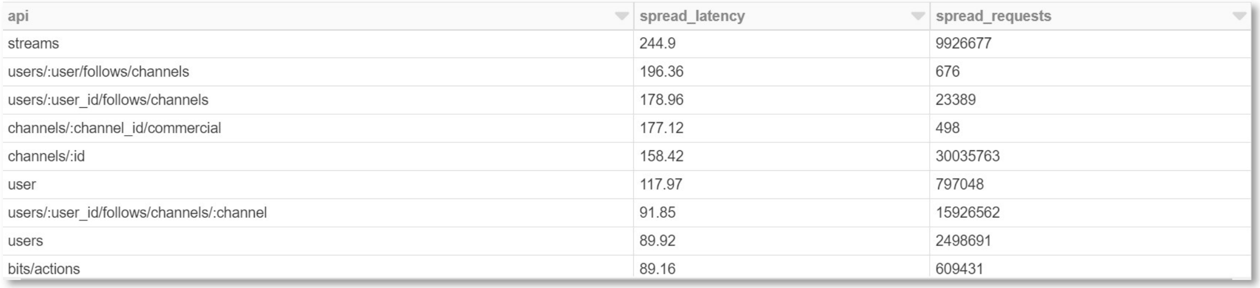
NEXT STEPS

- **Predict the Web Traffic to Further Improve the API Performance**

If we choose either **team_emote** or **one_more_thing**, there are still potential for better response speed, especially if the web traffic increases. Therefore, we can run a time series model to forecast the web traffic. If the traffic will go beyond a certain threshold in the short future, we should either accelerate the improvement of API performance or consider partnering with **the_ogp**.

- **More Efficient Developer Management**

Currently, multiple developers design the same APIs at the same time, but the performance varies a lot among different authors. In the future, we can manage the partnership with APIs more efficiently, especially for those whose spread of latency and total requests are both large as listed below. For those APIs, we can go through each API one by one to choose the developer that fits our requirement the best, either with the smallest latency or with the highest requests.



api	spread_latency	spread_requests
streams	244.9	9926677
users/:user/follows/channels	196.36	676
users/:user_id/follows/channels	178.96	23389
channels/:channel_id/commercial	177.12	498
channels/:id	158.42	30035763
user	117.97	797048
users/:user_id/follows/channels/:channel	91.85	15926562
users	89.92	2498691
bits/actions	89.16	609431

- **Better Monetization of Certain Functions**

Referred back to the top 10 most active APIs, **“bits/actions”** was ranked the tenth most frequently used. That being said, users are actively involved with the donation and interaction with the creator. We can improve our current monetization policy by identifying the hourly, weekly or monthly pattern (I have already cleaned the data for further analysis), so we can devise the corresponding promotion strategies.

** For the technical parts (SQL queries and Python work), please refer to **CindyWang__API__DataCleaning (HTML)** and **CindyWang__API__DataExploration (HTML)** in the zip file.