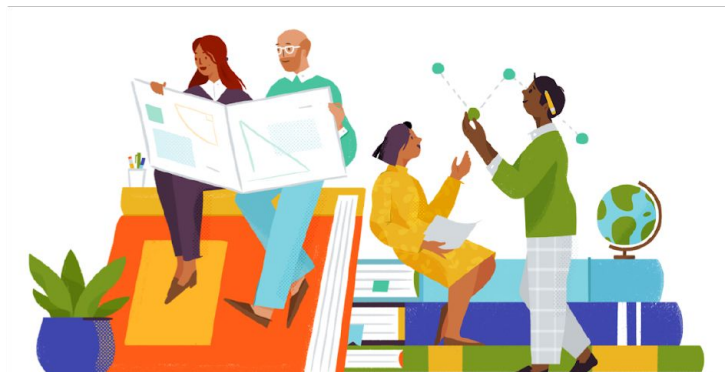


Search at Slack

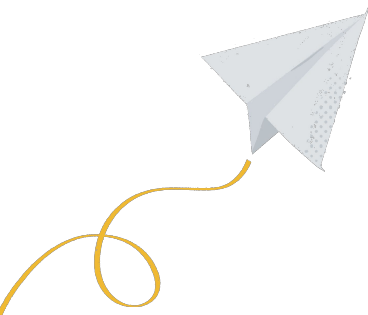
#search



About Me

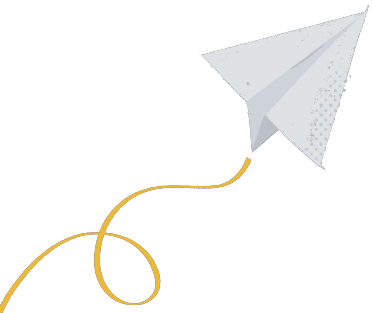
Search & Discovery Engineer @ Slack

Formerly Search @ Lucidworks

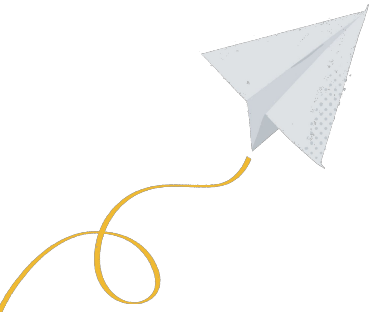


Agenda

1. Introduction
2. Slack Architecture
3. Indexing Architecture
4. EKM
5. I18n
6. Query Architecture



Introduction



Slack Stats

10M

Daily Active
Users

85k

Paying Customers
on Slack

65

Fortune 100
Customers

150

Expanded to 150+
Countries

1500

Apps in the Slack
Directory

Slack brings together people, data and applications




Joolep Inc.
 Rose Nguyen

🔍 world cup ad in: #team-marketing from: @Michael Chen from: @Paige Robinson Clear X


Messages Files


21 Results Most recent Filter by

team-marketing - Jun 14


 **Paige Robinson** 11:23 AM
Our world cup ad ran after the first goal was scored!
Here's a link to the video: <http://www.youtube.com...>

team-marketing - Jun 11


 **Michael Chen** 4:07 PM
Our world cup ads begin airing this week! We'll be featured during these games:

 **World Cup Ad Schedule**
Nami Taylor - June 1st, 2018

team-marketing - May 18




 **Paige Robinson** 1:50 PM
Hi team! We have the final edits for our world cup ad series. Join us at 3pm for a screening party.

team-marketing - April 2

 **Paige Robinson** 10:27 AM
Some exciting news... we just signed off on the

Filter by

People

-  Paige Robinson
-  Michael Chen
-  Mani Sung

More...

Channels & Direct Messages

- # team-marketing
- # videos-world-cup
- # brainstorming

More...

Date

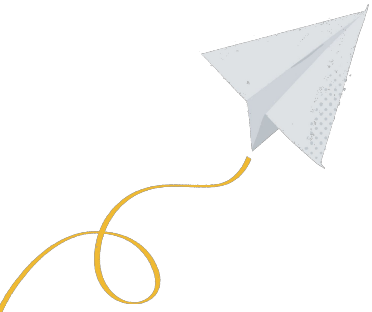
Starting

Ending

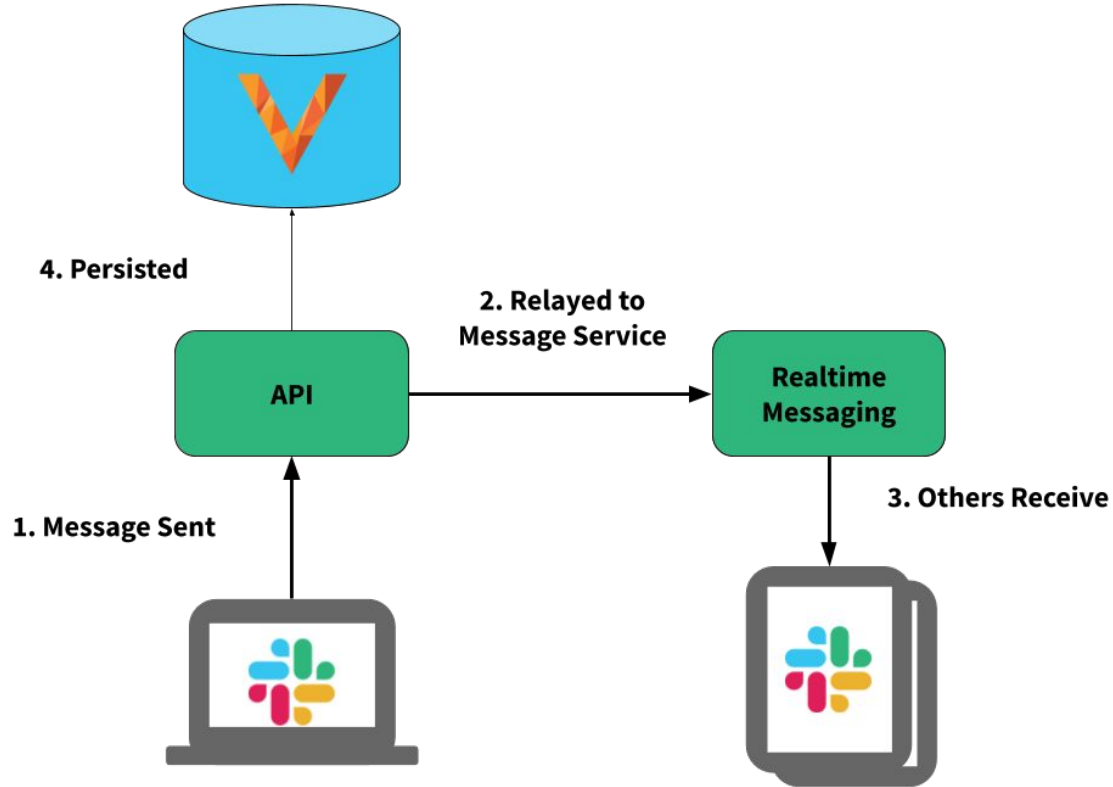
Include results from:

- Apps and bots
- Channels I'm not a member of

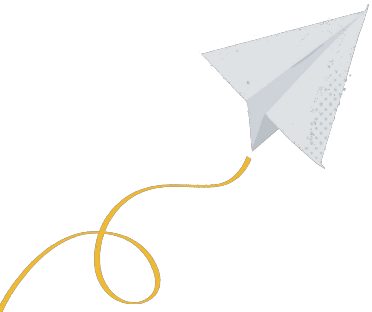
Slack Architecture



Slack Architecture



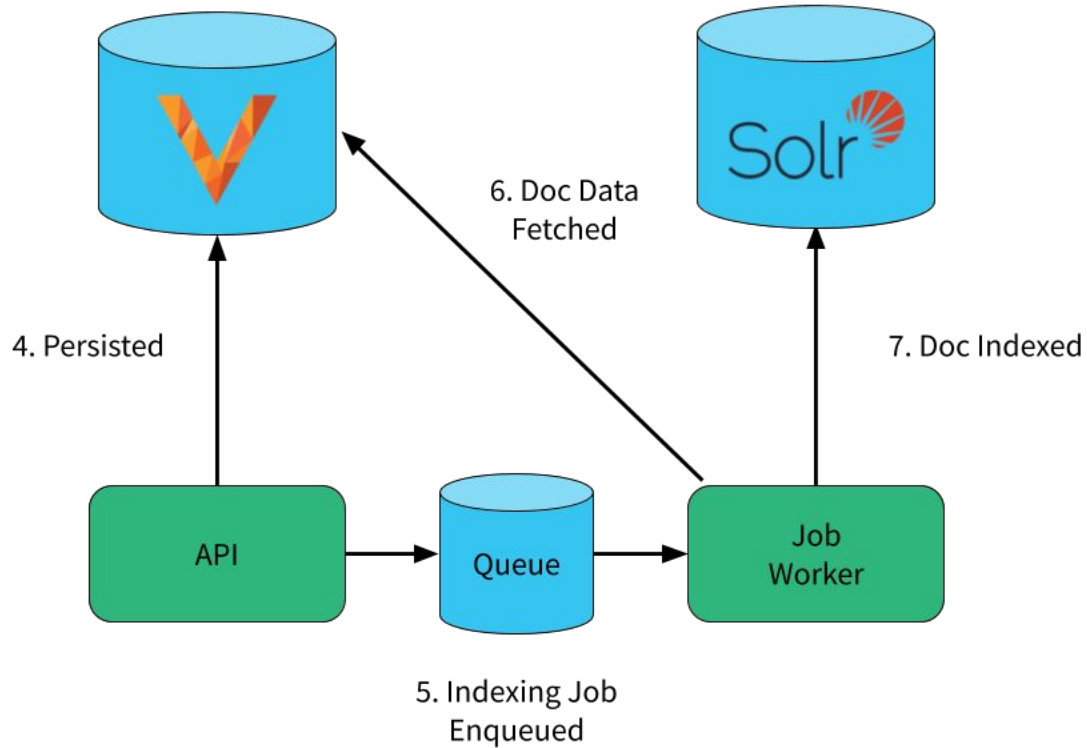
Search At Slack



Search Collections

name	last timestamp_tl (excl)	stripes	shards/stripes
chat-computed_19700101_20190602_v0			
live_20190602			
live_20190603			
live_20190604			
live_20190605			
live_20190606			
live_20190607			
live_20190608			

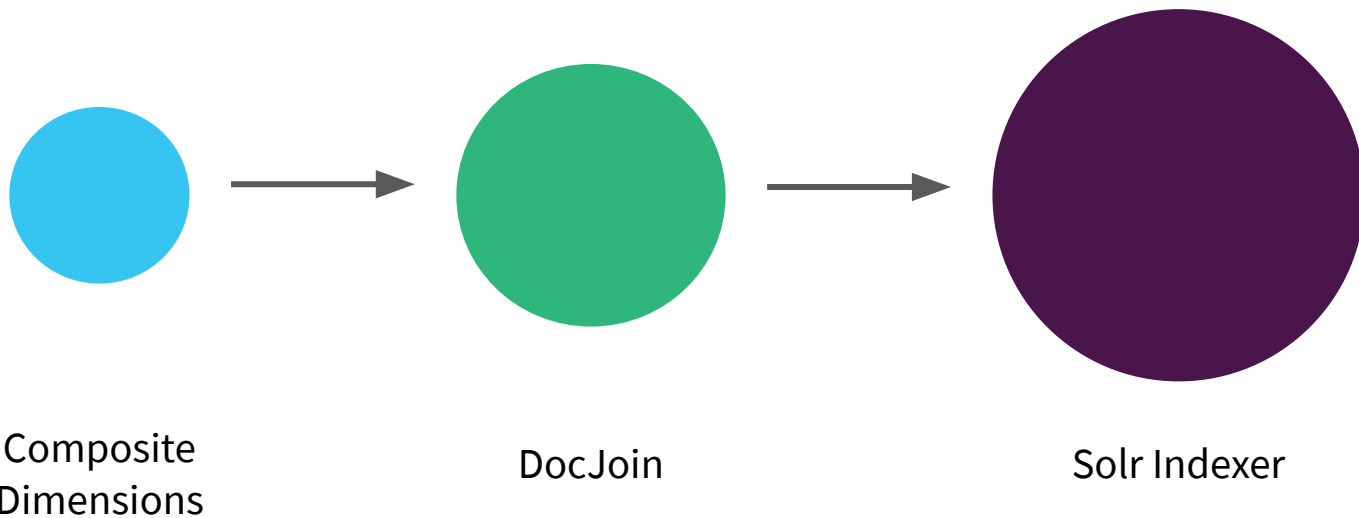
Indexing



Build An Offline Indexing Pipeline

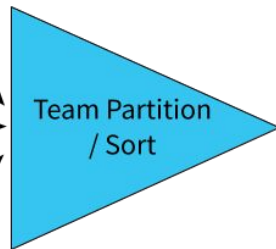
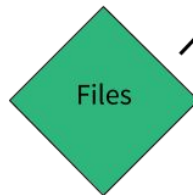
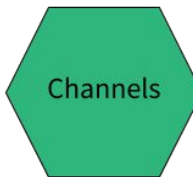
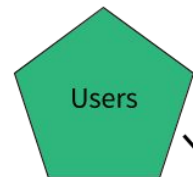


The Offline Indexing Pipeline

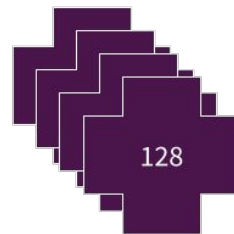


Composite Dimensions

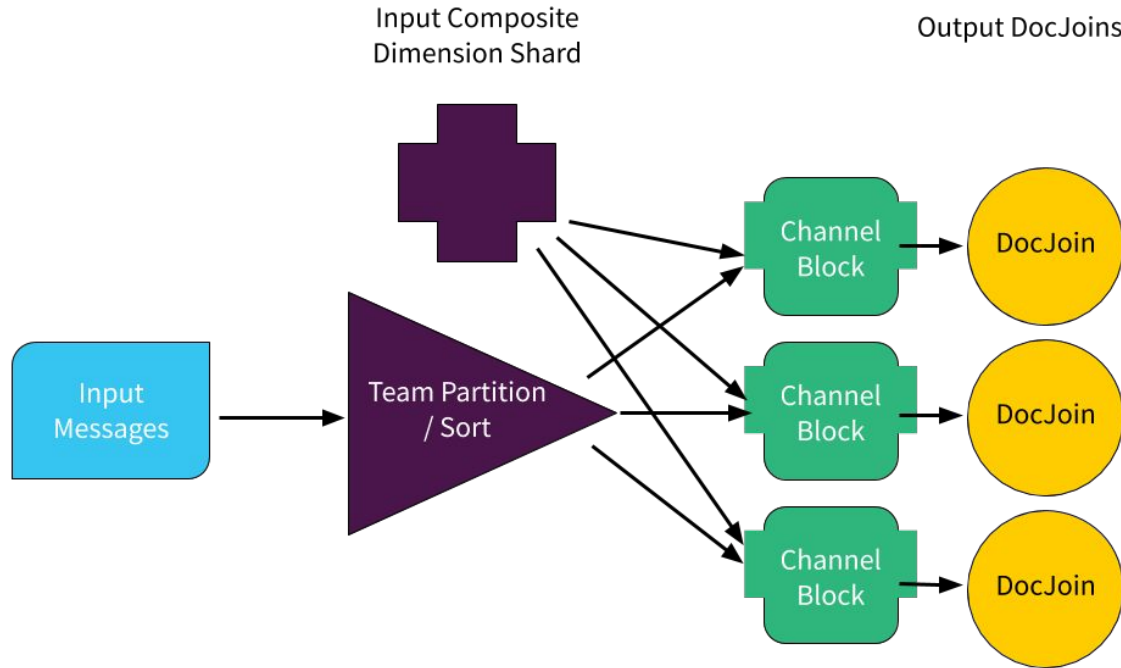
Input Dimension
Tables



Output Composite
Dimension Shards



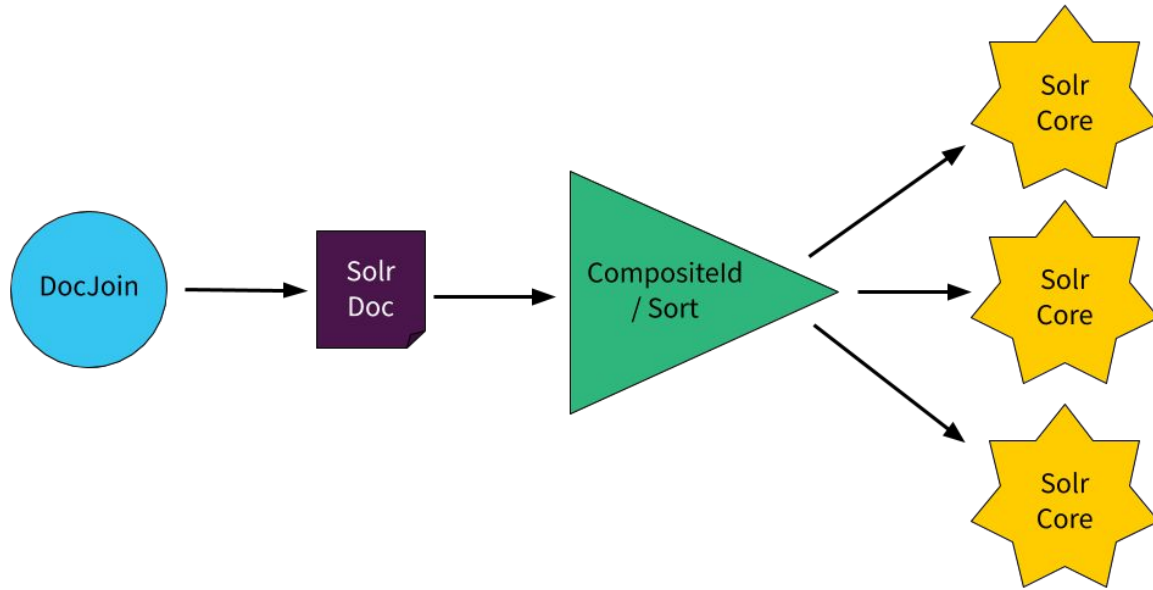
DocJoin



Solr Indexer

Input DocJoin

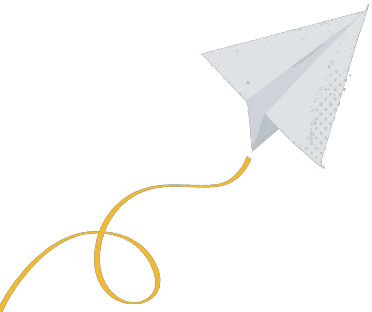
Output Solr Shards






Search Collections

name	last timestamp_tl (excl)	stripes	shards/stripes
chat-computed_19700101_20190602_v0			
live_20190602			
live_20190603			
live_20190604			
live_20190605			
live_20190606			
live_20190607			
live_20190608			

EKM



Enterprise Key Management

-  Customers control the key used to encrypt their data
-  An EKM customer's search index lives on an encrypted filesystem
-  Each EKM customer's search index is also built separately from
everyone else's

<https://slack.engineering/engineering-dive-into-slack-enterprise-key-management-1fce471b178c>

l10n and i18n

10 letters

Localization (l10n)

n. The process of adapting a product and its content to meet the requirements of a specific language or locale.

18 letters

Internationalization (i18n)

n. The process of designing and building an application so that it can be adapted to support various languages & locales

What Languages & Locales does Slack support?

1. German
2. English (UK)
3. English (US)
4. Spanish (Spain)
5. Spanish (Latin America)
6. French (France)
7. Portuguese (Brazil)
8. Japanese

How do we decide what language to tokenize in?

Workspace Language Preferences

Workspace Language Close

Set the language for your workspace. This affects system notifications, Slackbot messages, and sign-up emails. Your workspace language is currently set to: **English (US)**

English (US) ▼

Save

User Level Language Preferences

Your preferences for The Resistance


Language & Region


Language

English (US) ▼

Choose the language you'd like to use with Slack.

Example:

 **Han Solo** 8:59 AM
La gente cuenta con nosotros. ¡La galaxia cuenta con nosotros!

 1

Workspace Preference: **English**

[la, gent, cuenta, con, nosotros, la, galaxia, cuenta, con, nosotros]

User Preference: Spanish

[gent, cuent, galaxi, cuent]

Example:



Han Solo 8:59 AM

La gente cuenta con nosotros. ¡La galaxia cuenta con nosotros!



1

Workspace Preference: English

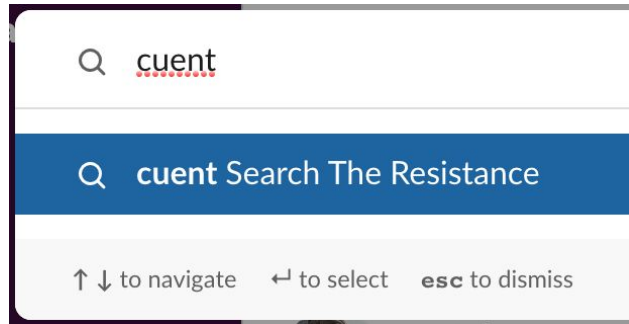
[la, gent, **cuenta**, con, nosotros, la,
galaxia,
cuenta, con, nosotros]

User Preference: Spanish

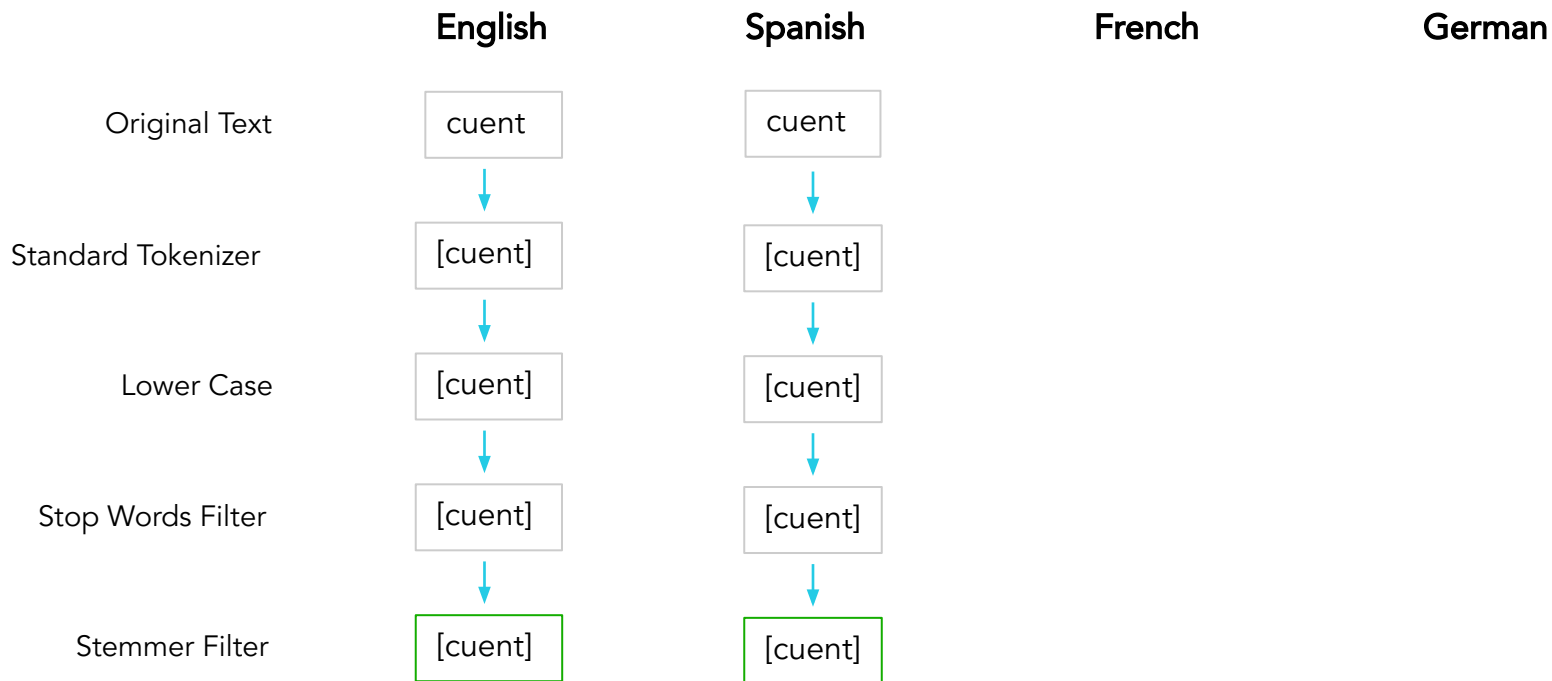
[gent, **cuent**, galaxi,
cuent]

```
{  
  "id": "chat_1001794-105-1559770571000400",  
  "type": "chat",  
  "team_id_t1": 1001794,  
  "user_id_t1": 112,  
  "user_s": "hansolo",  
  "channel_ids_t1m": [105],  
  "chan_type_s": "C",  
  "timestamp_t1": 1559770571000400,  
  "msg_t": "La gente cuenta con nosotros. ¡La galaxia cuenta con nosotros!",  
  "msg_es": "La gente cuenta con nosotros. ¡La galaxia cuenta con nosotros!",  
  "reactions_t1": 1,  
  "_version_": 1,  
  "reactions_tokens_tm_as_is": ["rebel 🗡️113"]}]  
}
```

What happens now when we search?



Tokenize the Query in ALL 8 Supported Languages



Is there a match?

Query

Message

msg_t (English)

[cuent]

[la, gent, **cuenta**, con, nosotros, la, galaxia, cuenta, con, nosotros]

msg_es (Spanish)

[cuent]

[gent, **cuent**, galaxi, cuent]

Is there a match?

Query

Message

msg_t (English)

[cuent]

[la, gent, cuenta, con, nosotros, la, galaxia, cuenta, con, nosotros]

msg_es (Spanish)

[cuent]

[gent, cuent, galaxi, cuent]




Search Result


Q cuent

Messages Files Channels

2 results Most recent ▾

millennium-falcon - Jun 6th

 **Han Solo** 11:21 AM
La gente **cuenta** con nosotros. ¡La galaxia **cuenta** con nosotros!

 1 reaction

Potential Improvements

- Being able to detect the language at index time in addition to relying on the team and user locale settings
- Detecting the language at query time so that we can search only on specific language fields
- Smarter Tokenization - Using one field to index all languages . We need to search on one field only

CJK

Indo-European language examples: English, Spanish, Portuguese, French, German
CJK - Chinese, Japanese, and Korean languages



“RAWRGWAWGGR.” - Chewbacca
Shyriiwook



Galactic Basic

How do we support CJK?

ソフトウェアは日本で人気があります



Contains CJK [Chinese, Japanese, Korean] characters?



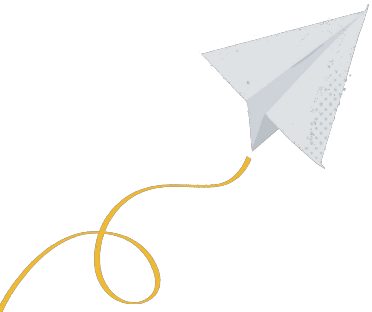
msg_ja

text	ソフトウェア	日本	人気
raw_bytes	[e3 82 bd e3 83 95 e3 83 88 e3 82 a6 e3 82 a7 e3 82 a2]	[e6 97 a5 e6 9c ac]	[e4 ba ba e6 b0 97]
start	0	7	10
end	6	9	12

msg_cjk3

text	ソフト	フトウ	トウェ	ウェア	エアは	アは日	は日本	日本で
raw_bytes	[e3 82 bd e3 83 95 e3 83 88]	[e3 83 95 e3 83 88 e3 82 a6]	[e3 83 88 e3 82 a6 e3 82 a7]	[e3 82 a6 e3 82 a7 e3 82 a2]	[e3 82 a7 e3 82 a2 e3 81 af]	[e3 82 a2 e3 81 af e6 97 a5]	[e3 81 af e6 97 a5 e6 9c ac]	[e6 97 a5 e6 9c ac]
start	0	1	2	3	4	5	6	7
end	3	4	5	6	7	8	9	10

Query Architecture

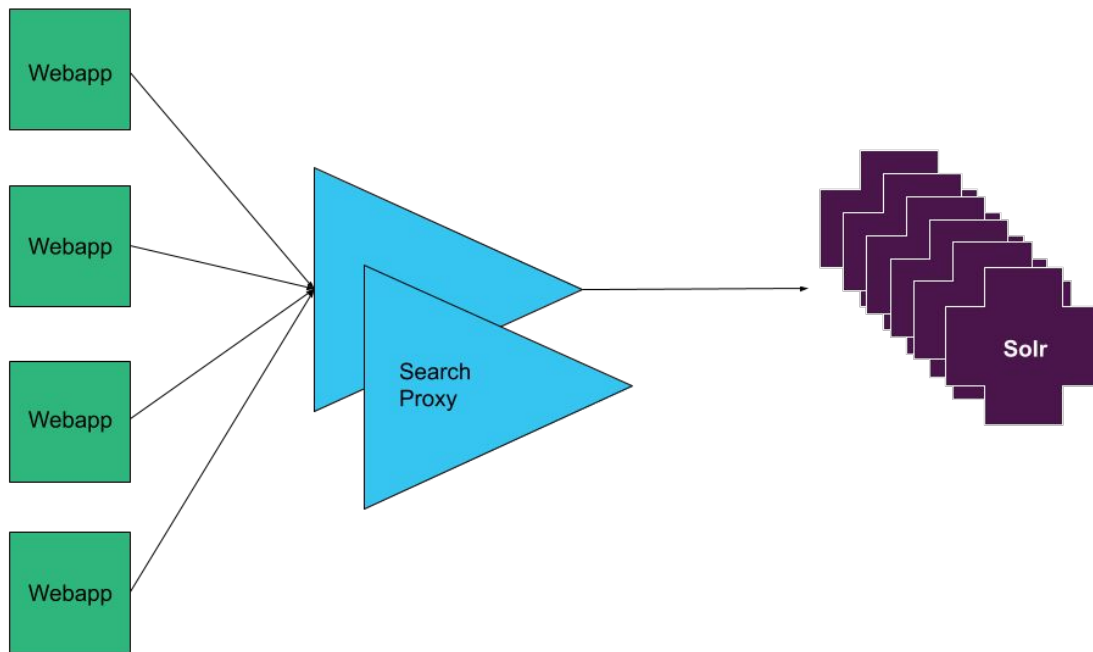


Query Layer

Slack API Service

Search Proxy

Solr Cluster



Joolep Inc.
 Rose Nguyen

world cup ad in: #team-marketing from: @Michael Chen from: @Paige Robinson Clear X

Messages Files

21 Results Most recent Filter by

team-marketing - Jun 14

Paige Robinson 11:23 AM
Our world cup ad ran after the first goal was scored!
Here's a link to the video: <http://www.youtube.com...>

team-marketing - Jun 11

Michael Chen 4:07 PM
Our world cup ads begin airing this week! We'll be featured during these games:

World Cup Ad Schedule
Nami Taylor - June 1st, 2018

team-marketing - May 18

Paige Robinson 1:50 PM
Hi team! We have the final edits for our world cup ad series. Join us at 3pm for a screening party.

team-marketing - April 2

Paige Robinson 10:27 AM
Some exciting news... we just signed off on the

Filter by

People

- Paige Robinson
- Michael Chen
- Mani Sung

More...

Channels & Direct Messages

- # team-marketing
- # videos-world-cup
- # brainstorming

More...

Date

Starting

Ending

Include results from:

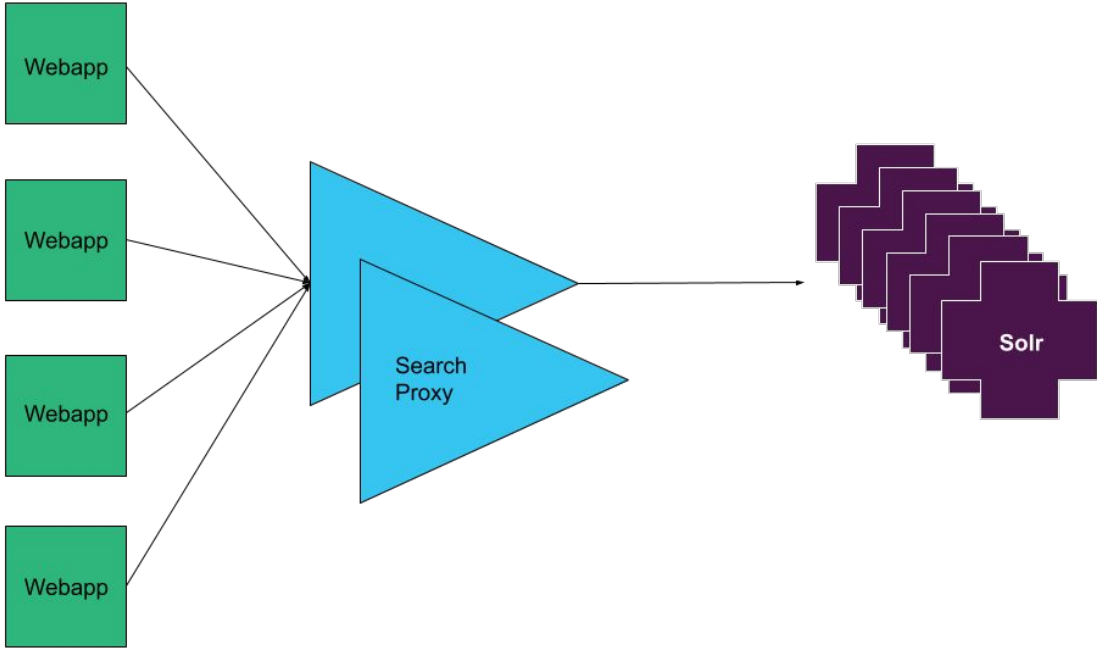
- Apps and bots
- Channels I'm not a member of

Proxy Service

Slack API Service

Search Proxy

Solr Cluster





Thank You!

We're hiring! slack.com/jobs

