

從開源走向世界 從台灣影響國際

聊聊為何開源值得耽誤你的時間

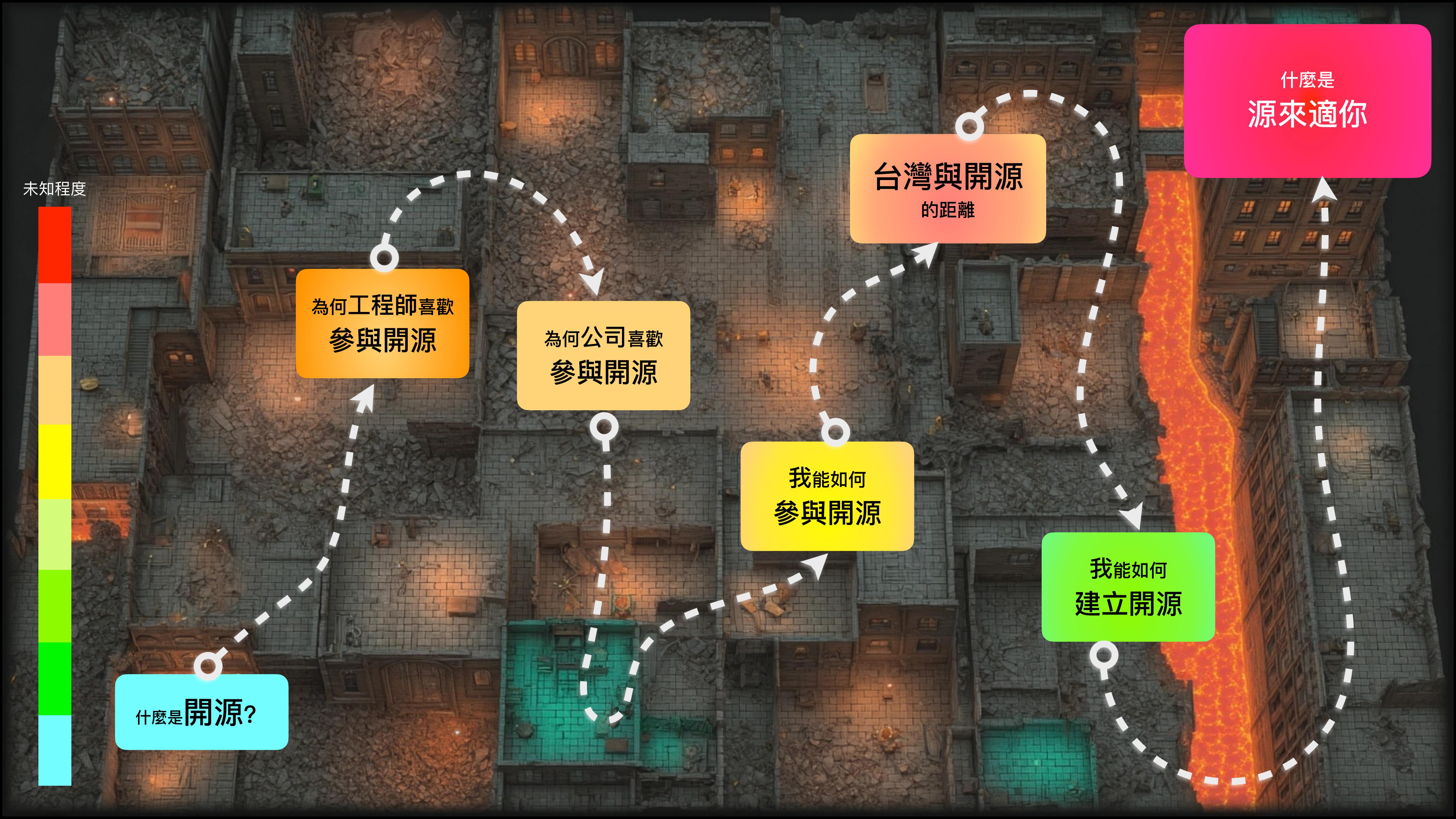
Chia-Ping Tsai

蔡嘉平

www.chia7712.tw

- Member @ Apache
- 創辦人 @ 源來適你
- 資訊工程博士 @ 成功大學
- PMC member @ Apache Kafka
- PMC member @ Apache HBase
- PMC member @ Apache YuniKorn
- PC member @ HBaseConAsia
- PC member @ Kafka Summit





什麼是
源來適你

未知程度

什麼是
開源?

為何工程師喜歡
參與開源

為何公司喜歡
參與開源

台灣與開源
的距離

我能如何
參與開源

我能如何
建立開源

什麼是開源？

什麼是開源?
什麼是參與開源?

我把程式碼放到 Github 上公開



100分

我發送 patch 到別人的開源專案



100分

我使用開源軟體打造核心服務？



100分

感謝大家的聆聽
請款單在哪裡簽？

我把程式碼放到
Github 上公開

有社群嗎?
有使用者嗎?
有影響力嗎?

沒有? 那只是自嗨而已

我發送 patch 到
別人的開源專案

是有影響力的貢獻嗎?
該專案是有潛力的嗎?

沒有? 沒關係有貢獻就很好

我使用開源軟體
打造核心服務?

有了解原理嗎?
出問題能修復嗎?
掌握技術發展方向嗎?

沒有? 那就是數位殖民

[Pull requests 257](#)[Actions](#)[Projects 1](#)[Security 1](#)[Insights](#)

kafka

Public

[Edit Pins](#)[Unwatch 1056](#)[Fork 14.4k](#)[Starred 30.1k](#)[trunk](#)[79 Branches](#)[269 Tags](#)[Go to file](#)[t](#)[Add file](#)[Code](#)[About](#)**lorcanj** and **Lorcan** KAFKA-4650: Add unit tests for GraphNode class (#...)

e88c10d · 6 hours ago

15,767 Commits

[.github](#)

MINOR: Update the supported tags in docker_scan.yml (...)

2 days ago

[bin](#)

MINOR: Supress stdout when checking Log4j 1.x configu...

last month

[checkstyle](#)

MINOR: Move logDirs config out of KafkaConfig (#19579)

5 days ago

[clients](#)

KAFKA-18783 : Extend InvalidConfigurationException rel...

19 hours ago

[committer-tools](#)

KAFKA-18942: Add reviewers to PR body with committer-...

2 months ago

[config](#)

KAFKA-14485: Move LogCleaner to storage module (#19...

last month

[connect](#)

MINOR: replace .stream().forEach() with .forEach() (#196...

2 weeks ago

[coordinator-common/src](#)

KAFKA-19216: Eliminate flakiness in kafka.server.share.S...

2 weeks ago

[core](#)

MINOR: Cleanup Core Module- Scala Modules (1/n) (#19...

10 hours ago

[docker](#)

MINOR: bump zstd minor version to 1.5.6-10 (#19512)

last month

[docs](#)

KAFKA-19287 document all group coordinator metrics (#...

7 hours ago

[examples](#)

MINOR: Add retry mechanism to EOS example (#15561)

last year

[generator/src](#)

KAFKA-18955: Fix infinite loop and standardize options in...

2 months ago

[gradle](#)

MINOR: update the gradle version for java 24 support (#...

11 hours ago

[group-coordinator](#)

KAFKA-17747: [3/N] Get rid of TopicMetadata in Subscrib...

2 days ago

[jmh-benchmarks](#)

KAFKA-19302 Move ReplicaState and Replica to server m...

2 days ago

[licenses](#)

KAFKA-17593; [1/N] Introduce re2j dependency (#17634)

7 months ago

有社群嗎？

Mirror of Apache Kafka

scala

kafka

[Readme](#)[Apache-2.0, Apache-2.0 licenses found](#)[Code of conduct](#)[Security policy](#)[Activity](#)[Custom properties](#)[30.1k stars](#)[1.1k watching](#)[14.4k forks](#)[Report repository](#)

Releases

[269 tags](#)[Create a new release](#)

Packages

No packages published

[Publish your first package](#)

Contributors 1,256



POWERED BY

Apache Kafka is the most popular open-source stream-processing software for collecting, processing, storing, and analyzing data at scale. Most known for its excellent performance, low latency, fault tolerance, and high throughput, it's capable of handling thousands of messages per second. With over 1,000 Kafka use cases and counting, some common benefits are building data pipelines, leveraging real-time data streams, enabling operational metrics, and data integration across countless sources.

Today, Kafka is used by thousands of companies including over 80% of the Fortune 100. Among these are Box, Goldman Sachs, Target, Cisco, Intuit, and more. As the trusted tool for empowering and innovating companies, Kafka allows organizations to modernize their data strategies with event streaming architecture. Learn how Kafka is used by organizations in every industry - from computer software, financial services, and health care, to government and transportation.

AddThis

Apache Kafka is used at AddThis to collect events generated by our data network and broker that data to our analytics clusters and real-time web analytics platform.

aven

Aiven is a cloud platform for open source technologies. We provide Apache Kafka as a managed service on public clouds, and use it internally to run and monitor our platform of tens of thousands of clusters.

Ancestry

Kafka is used as the event log processing pipeline for delivering better personalized product and service to our customers.

Appsflyer

Apache Kafka is a core piece of Appsflyer large-scale event driven architecture.

It enables Appsflyer to run a data pipeline that streams tens of billions of events on a daily basis

AXON

Axon uses Apache Kafka to power mission critical infrastructure that enables first responders do their jobs effectively.

BARCLAYS

Barclays utilizes Kafka for streaming and analytical information.

bunq

Apache Kafka aggregates high-flow message streams into a unified distributed pubsub service, brokering the data for other internal systems as part of Boundary's real-time network analytics infrastructure.

CISCO

Cisco is using Kafka as part of their OpenSOC (Security Operations Center). More details [here](#).

CLOUDFLARE

Cloudflare uses Kafka for our log processing and analytics pipeline, collecting hundreds of billions of events/day and data from

adidas

adidas uses Kafka as the core of Fast Data Streaming Platform; integrating source systems and enabling teams to implement real-time event processing for monitoring, analytics and reporting solutions.

agoda

Apache Kafka powers the backbone of Agoda's data pipeline with trillions of events streaming through daily across multiple data centers. The majority of the events are destined for analytical systems and directly influence business decisions at one of the world's fastest growing online travel booking platforms.

airbnb

Used in our event pipeline, exception tracking & more to come.

allegrograph

AllegroGraph and Kafka are used together as an Entity Event Knowledge Graph platform in diverse settings such as call centers, hospitals, insurance companies, aviation organizations and financial firms. By coupling AllegroGraph with Kafka, users can create a real-time decision engine that produces real-time event streams based on computations that trigger specific actions.

amadeus

Apache Kafka is used for both real-time and batch data processing, and is the chosen event log technology for Amadeus microservice-based streaming applications. Kafka is also used for operational use cases such as application logs collection.

altair

Altair Panopticon(TM) incorporates Kafka into its stream processing engine. The software allows business users to build stream processing applications that subscribe to streaming data inputs, including Kafka topics and other real-time message queues, retrieve from SQL, NoSQL and time series databases, join data streams and tables, aggregate streams within defined time windows, and confinate, filter, and merge data streams. The system can output processed data flows to Kafka or email, or write to databases like Kudu+, InfluxDB, or any SQL database.

DATA SIFT

Apache Kafka is used at DataSift as a collector of monitoring events and to track user's consumption of data streams in real time. [DataSift architecture](#)

ANTS

Ants.vn uses Kafka in production for stream processing and log transfer (over 5B messages/month and growing)

ATRUVIA

At Atruvia we use Apache Kafka to share events within the modern banking platform.

bandwidth

Bandwidth uses Kafka in their Communications Platform to process call and message data records for billing and data analytics. Kafka also powers Bandwidth application metrics, system metrics and log aggregation pipelines.

jack henry digital

The Banno Digital Platform from Jack Henry enables community financial institutions to provide world-class service in today's digital landscape. The Banno team integrates various streams of data through Apache Kafka, reacting to events as they occur, to provide innovative banking solutions.

AXIOS

The Data Platform team at Axios uses Kafka to stream real-time data from our products into our data lake where we are able to run exploratory queries to understand our data and build machine learning models. Kafka enables us to automate the collection of data from various parts of the organization for ingestion into our data platform.

EMERGING THREATS

Emerging threats uses Kafka in our event pipeline to process billions of malware events for search indices, alerting systems, etc.

eXponential

Eponential is using Kafka in production to power the events ingestion pipeline for real time analytics and log feed consumption.

BookingSync

BookingSync is using Kafka for its own implementation, and also empowers its users to store Events defined in their Event Models into Kafka.

FOURSQUARE

We use Kafka to collect all metrics and events generated by the users of the website.

globeo

Kafka is used at Globo to queue and stream real-time and batch data to our Big Data event ingestion. It is also used for our logging systems.

Bytedance

ByteDance uses Kafka as a data hub to collect events and logs for a variety of services such as online model training, stream data processing and real-time data analytics.

Cerner

Kafka is used with HBase and Storm as described [here](#).

Cityzen Data

Citizen Data uses Kafka as well, we provide a platform for collecting, storing and analyzing machine data.

affiliate

Apache Kafka is used at CJ Affiliate to process many of the key events driving our core product. Nearly every aspect of CJ's products and services currently benefit from the speed and stability this provides; additionally

CloudFlare

Kafka is powering our high-flow event pipeline that ingests over 1.2 billion

CloudPhysics

HackerRank uses Kafka as events as a service platform. We publish all the internal activity on our infrastructure into Kafka, and a wide range of internal services subscribe to it.

CloudIO

Kafka powers CloudIO Flow which is an enterprise grade data ingestion platform that streams data from multiple sources in

Grab

Graylog2 is a free and open source log management and data analysis system. It's using Kafka as default transport for Graylog2 Radio. The use case is described [here](#).

CloudPHYSICS

Kafka is powering our high-flow event pipeline that ingests over 1.2 billion

nuuly

We use Kafka in production for online and near real-time solutions. Kafka is a core part for many products, such as our Credit Card System.

covage

Covage is an infrastructure provider, designing, deploying and operating high speed open access networks. At the very heart of our IT platform, Kafka is ensuring propagating our business workflows' events among all applications.

CRITEO

Criteo takes advantage of Apache Kafka at the heart of the core business. Kafka powers our business log collection pipeline and streaming infrastructure. We have tens of Kafka clusters deployed over multiple data centers across three continents processing up to 30 million messages/sec.

dash0

Dash0 leverages Apache Kafka as its core event streaming platform to ingest, process, and distribute customer OpenTelemetry data at scale. Kafka's distributed architecture allows Dash0 to handle high volumes of real-time telemetry data with low latency.

optimove

Optimove powers its real-time event streaming with Confluent Cloud and Apache Kafka to maintain always-up-to-date views of customer data and support machine learning algorithms that maximize campaign effectiveness.

ORACLE

Oracle provides native connectivity to Kafka from its Enterprise Service Bus product called OSB (Oracle Service Bus) which allows developers to leverage OSB built-in mediation capabilities to implement staged data pipelines.

instaclustr

Instaclustr provides a fully managed and hosted Apache Kafka service. We use Kafka internally for our streaming Change Data Capture framework used by LiveDesign, our enterprise informatics solution.

JitBit

JitBit powers Jitbit's logging pipeline, analytics events, but most importantly, queues full-text search indexing for hundreds of millions of help desk tickets in our cloud system.

OPT

OPT.nc uses Kafka to process large amount of logs, to queue sms messages sending/ queuing, spread and share internal data with Streams across various telecommunications like Telecommunication, Financial services, Geographical Information System (GIS) and Post Office delivery process.

OREXES

Deephaven is a query engine for streaming workloads. Deephaven enables you to ingest and transform Kafka feeds as live dataframes.

DEVSISTERS

We use Apache Kafka (and Kafka Streams) to collect and ingest all of our game service logs (including analytics, servers, or access logs). Apache Kafka has been one of the core components of our data pipeline from early 2015.

DREAMII

We use Apache Kafka heavily for data ingestion to the Data platform, streaming as well as batch analytics, and for our microservices to communicate with one another. Kafka is a core component of the overall Dream11 Tech stack.

EDENLAB

Edenlab FHIR engineers use Kafka to manage real-time data flow between dozens of microservices in our product - [Kotlin Interoperability Suite](#) - a low-code FHIR-based infrastructure for healthcare data management (contains FHIR Server, Terminology Service, and Mapper). Edenlab is a custom software and product development company focused primarily on healthcare data interoperability.

DEVCOM

Customer relationship management, order management system, incident management system, and tech support system.

OHV.com

OVH uses Kafka in production for over a year now using it for event bus, data pipeline for antiddos and more to come.

PayPal

At PayPal, Kafka is used for first-party tracking, application health metrics streaming and aggregation, database synchronization, application log aggregation, batch processing, risk detection and management, and analytics and compliance, with each of these use-cases processing over 100 billion messages per day. See [this](#).

PARSE.LY

Kafka is used for all data integration of analytics event data.

LinkSmart

Kafka is used at LinkSmart as an event stream feeding Hadoop and custom real time systems.

QUIKEY

At Quixey, the Search Engine for Apps, Kafka is an integral part of our eventing, logging and messaging infrastructure.

QUODOSOFT

At Quodsoft, as part of the bigger tech network organization behind Germany based KLINGEL group, we build a big scale e-commerce platform called Next Level Commerce (NLC). NLC is based on the principle of customer-oriented verticalization which allows us maximum autonomy for our teams. With our microservices architecture and compliance, with each of these use-cases processing over 100 billion messages per day. See [this](#).

EXOSCALE

Exoscale uses Kafka in production.

Recursion

Recursion uses Kafka Streams to power its data pipeline for drug discovery efforts. Kafka is used to coordinate various services across the company. For more information about the use case see [this Kafka Summit talk](#).

Reddit

Reddit deploys Kafka to power our product, infrastructure and analytics. Apache Kafka is flexible and stable, allowing Reddit communities to thrive while our engineers build with confidence.

Rabobank

Rabobank is one of the 3 largest banks in the Netherlands. Its digital nervous system, the Business Event Bus, is powered by Apache Kafka. It is used by an increasing amount of financial processes and services, one of which is Rabo Alerts. This service alerts customers in real-time upon financial events and is built using Kafka Streams.

mate1

Mate1 kafka is used at Mate1 as our main event bus that powers our news and activity feeds, automated review systems, and will soon power real time notifications and log distribution.

metamarkets

As an IOT company, we use Kafka as the backbone of our event streaming/ telemetry collection pipeline and as an event bus for microservices.

moengage

At MoEngage, Apache Kafka is one of the core components of our infrastructure and the backbone of all the events streaming services. We run over 25+ Kafka clusters with processing over 1 Million messages per second across all these clusters. find more about our Journey with Kafka so far at [Kafka Redesign and Lessons Learned](#).

Rollbar

Rollbar uses Kafka at large scale to store all incoming raw blobs. Kafka allowed us to have multiple workers and also allowed us to migrate to Kubernetes.

sematext

Sematext does kafka integration with our data processing pipeline.

SCHRÖDINGER

At Schrödinger, Kafka powers our physics-based computational platform by feeding data into our predictive modeling, data aggregation and in-memory data storage (OLAP cube creation) as well as our CEP/Alerts servers (see also: [SPM for Kafka](#))

hellosafe

Hellosafe is an international online insurance policy comparison platform. We publish all the internal activity on our infrastructure into Kafka, and a wide range of internal services subscribe to it. Both stateful and stateless

affiliate

HackerRank uses Kafka as events as a service platform. We publish all the internal activity on our infrastructure into Kafka, and a wide range of internal services subscribe to it.

CloudIO

CloudIO Flow which is an enterprise grade data ingestion platform. It's using Kafka as default transport for Graylog2 Radio. The use case is described [here](#).

CloudPHYSICS

Kafka is powering our high-flow event pipeline that ingests over 1.2 billion

Grab

Graylog2 is a free and open source log management and data analysis system. It's using Kafka as default transport for Graylog2 Radio. The use case is described [here](#).

CloudIO

Kafka is powering our high-flow event pipeline that ingests over 1.2 billion

De Facto Standard APIs: S3 and Kafka



Amazon S3

== de facto standard API for OBJECT STORAGE



有影響力嗎?



== de facto standard API for EVENT STREAMING



CONFLUENT



Confluent Cloud

CLOUDERA



Red Hat



Amazon MSK



Azure HDInsight

instaclustr



Azure Event Hubs



aiven



TIBCO



PULSAR

為何工程師喜歡參與開源？

你能從開源獲得的好處

- ① 提高實力
- ② 練習溝通
- ③ 認識大神
- ④ 獲取工作
- ⑤ 實現自我

源來適你 



“感謝開源讓我畢業直接來美國工作，然後認識各國為開源努力的大神”
蘇桓平 / Union.AI 

“參與開源，打造你的微斜槓職涯，建立工作之外的第二晉升機會”
楊博安 / Suse 

“我覺得開源讓小白能參與真的在業界被使用的世界級專案，被大神指導的同時還能建立人脈！”
邱譯 / Union.AI 

“通過參與開源可以學習別人的最佳實踐，拓寬視野，同時也能貢獻自己的力量做出改變”
曾冠博 / Suse 

“看到自己做得東西影響全世界覺得很開心”
陳楷訓 / Anyscale 

“透過參與社群，不僅能降低在貢獻開源專案時的技術與人脈門檻，同時許多前輩可供請益，從中吸取多元的思維方式與技術細節，進而大幅拓展自身的視野。”
黃竣陽 / Line 

“創造影響力，結識強者，提升眼界”
許秉倫 / LinkedIn 

“表面上參與開源最大的好處是讓我在沒有好的學歷下成為科技業 top 0.1% 但其實參與開源對我來說最大的好處就是認識這些人是真心熱愛寫程式，真心熱愛自己的東西被世界用到”
陳翰儒 / Union.AI 

“認識世界上最聰明的人，知道世界有多大，因為世界最先進最重要的技術都是開源的。最新最有趣的技術都是開源，做過幾個開源專案，也因此知道新創如何從零到一開發技術產品。”
莊偉赳 / Cloudera 

“參與開源全方位的提升了工程師的各式能力值。透過開源提升了在巨型 code base 裡 trace 的能力，並且接觸到了在平常較少機會碰到的複雜程式模型。”
紀登耀 / TapPay 

“做開源不僅能提升技術實力，相較於閉源貢獻也更透明，方便展示成果。還可拓展人脈，結識全球頂尖工程師，即使是 junior 也能接觸 principle engineer。”
劉奇聖 / Anyscale 

“俗話說的好「有人 code review 的孩子像個寶」。每每覺得自己程式寫的很糟簡直是一坨，看著開源前輩們溫柔的 code review，彷彿聽到他們說「傻孩子，你寫的程式，並沒有你想的那麼好」，砥礪我繼續精進”
李唯 / Astronomer 

工程師都需要「具移動性」的技能



「Even though the tech stack was scary in the beginning, I have learned so much. **Best part? Building new things from scratch or using open source teaches you skills you can take away with you.** When you join another company or start your own start-up, all the open-source libraries you know how to use are still available to use whenever you like.」

我自己延伸一下：矽谷公司也會自己 host 開源的服務，原因是

1. 避免重複造輪子
2. 利用社群的結果
3. 保有技術自主，避免被 商業模式 2 (Open Core) & 3 (SaaS) 綁定
4. 吸引人才，畢竟除了 Google 什麼都自己來以外，大家還是希望在工作累積的技能可以帶著走

由於這樣的原因，也需要聘請開源的人才維護/改進內部的服務。

清晨 5:46

源來適你

9月2日下午3:13 · 3月25日

時光飛逝三月就要結束，即將到來的不只是燥熱的四月，更有畢業和轉職的燥動。因此本週《科技開講》特別邀請台積電的Wei Chen副理，他將與大家分享豐富精彩的工作經驗：包含 Amazon, Meta 的八卦，從資料科學家、ML 工程師、AI 顧問、再到 Tech Manager 的心路歷程，以及他周遊台灣、日本、英國、美國各地的工作體驗！

歡迎3/30 週六早上10點一起來《科技開講》，會議室網址：<https://meet.google.com/uxt-ohwn-may>..... 查看更多

雲端大航海時代！

與 Google Cloud Taiwan 一起探索無限可能

主持人：蔡嘉平

TAM、TSE、軟AI/ML 工程師
你的下一站就在

I need JOB!

講者：黃介榮 / Google

九月 07 星期六 10AM

Jun Rao
Confluent 聯合創始人

4天 · 3月5日

「源來適你」是一個以開源為主的社群，從創立到現在總是圍繞在軟體上的開源。但其實開源無遠弗屆，在硬體的世界裡一樣有著很酷的開源概念！沒錯，這次的主題就是 RISC-V！

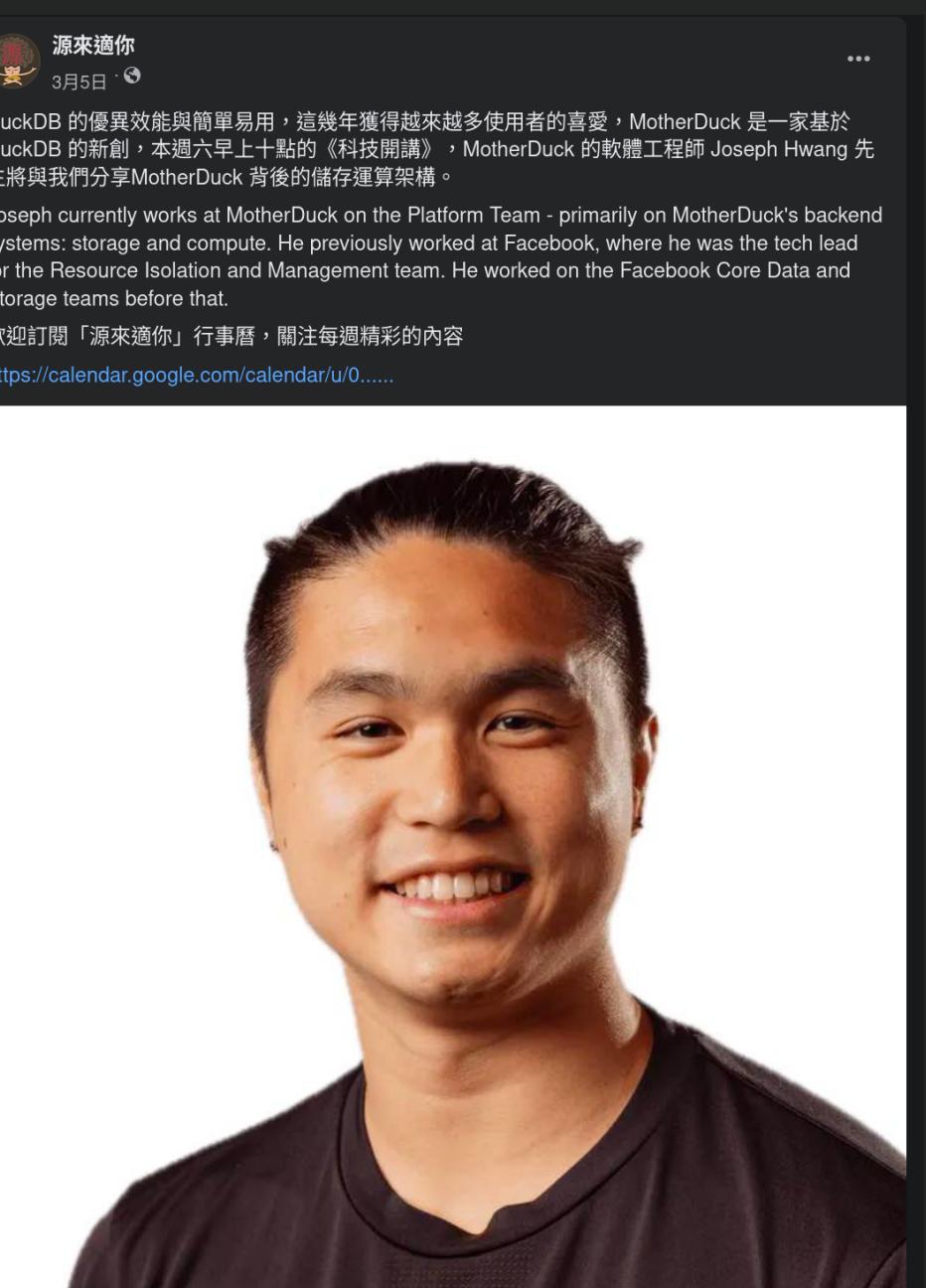
本週「科技開講」邀請晶心科技 楊凱名博士來跟大家暢談 RISC-V，也就是大家耳熟能詳的開源指令集架構（小編只講得出這句）。凱名博士在CPU設計有豐富的經驗，參與過許多RISC-V指令集的實做，在SoC整合也領導團隊進行各種系統偵錯和排除。

歡迎各位在本週 4/27 (六) 早上十點一起來暢談硬體上的開源！凱名博士會讓大家在會後變成 RISC-V 專家！..... 查看更多

Confluent 聯合創始人

源來適你 ● 追蹤中

中秋節月圓人更圓，在大家花式烤肉準備給醫生藥局做業績的時刻，Kafka 鯉魚戰隊也即將迎來 400 次的貢獻，為了紀念這個特別的日子與特別的數字，小編跪求來 confluent 聯合創始人 aka Kafka 大家長 aka 小編最崇拜的高手 - Jun Rao 來跟大家說幾句話，希望未來的日子我們繼續一起在開源上前進！



認識各路大神

科技開講

4月10日上午11:32 · 4月21日下午3:41

本週「科技開講」將由 Anyscale 的 Yifei 來為大家帶來一場知識饗宴！Yifei 作為TensorFlow的核心貢獻者，她不只是一位傑出的工程師，更在Anyscale 展現強大的領導能力帶領團隊克服許多技術挑戰！

如果想與 Yifei 請教任何技術問題，歡迎在本週六 4/13 早上十點加入會議室：<https://meet.google.com/uxt-ohwn-may>

或是訂閱「源來適你」行事曆不錯過任何一場精彩的活動：..... 查看更多

Yifei Feng

Engineering Manager @ Anyscale
Yifei leads the Infrastructure and SRE teams at @anyscale. Her teams focus on building a seamless, cost-effective, and scalable infrastructure for large-scale machine learning workloads.

TensorFlow top 20 contributors (>800 commits)
Before Anyscale, she worked on the open-source machine learning library TensorFlow for a few years at Google.

源來適你

8月21日下午3:41

經過上週熱熱鬧鬧的百人求職焦慮團，本週「科技開講」回歸傳統妥妥的技術(八卦)分享路線！

本週請來 Apple 鄭黃翔請來的 Nvidia 陳映嘉，雙主菜加值不加價，他們將聯手為各位帶來 AI 時代下一個兵家必爭之地：「AI在生醫領域的應用！」

不管你是腳在資訊還是腳在生醫，或是兩腳都在其他地方，跨領域的結合是職涯上不容易且高價值的事情，所以看熱鬧也好問八卦也不錯，歡迎各位朋友在 8/24 (六) 早上十點，一起來 <https://meet.google.com/uxt-ohwn-may> 共襄盛舉！

AI在生醫領域的應用

蛋白質結構與藥物設計、醫學影像

主持人：鄭黃翔 / Apple

小幫手：蔡嘉平博士

講者：陳映嘉 / NVIDIA

八月 24 星期六 10AM

源來適你

8月30日下午1:35 · 4月2日下午4:04

本週《科技開講》再度再度（很重要所以說三次）迎來大神工程師：蔡東邦 (DB Tsai) 任職於蘋果並擔任 Senior Engineering Manager，在團隊中負責多項熱門開源技術，從運算到資料處理、從 Spark (同時也是 PMC member), Flink, Hadoop 到 HBase。

蔡東邦最酷的事情是他不只有豐富的大規模資料處理經驗，也有在蘋果公司建立團隊的文化體驗！因此歡迎各位 4/6 (六) 早上十點一起來《科技開講》，讓我們一起站在大神的肩膀上！

會議室網址：<https://meet.google.com/uxt-ohwn-may>..... 查看更多

科技開講:刷機與人生 - 吳泓霖 (台灣 9/1 10AM)

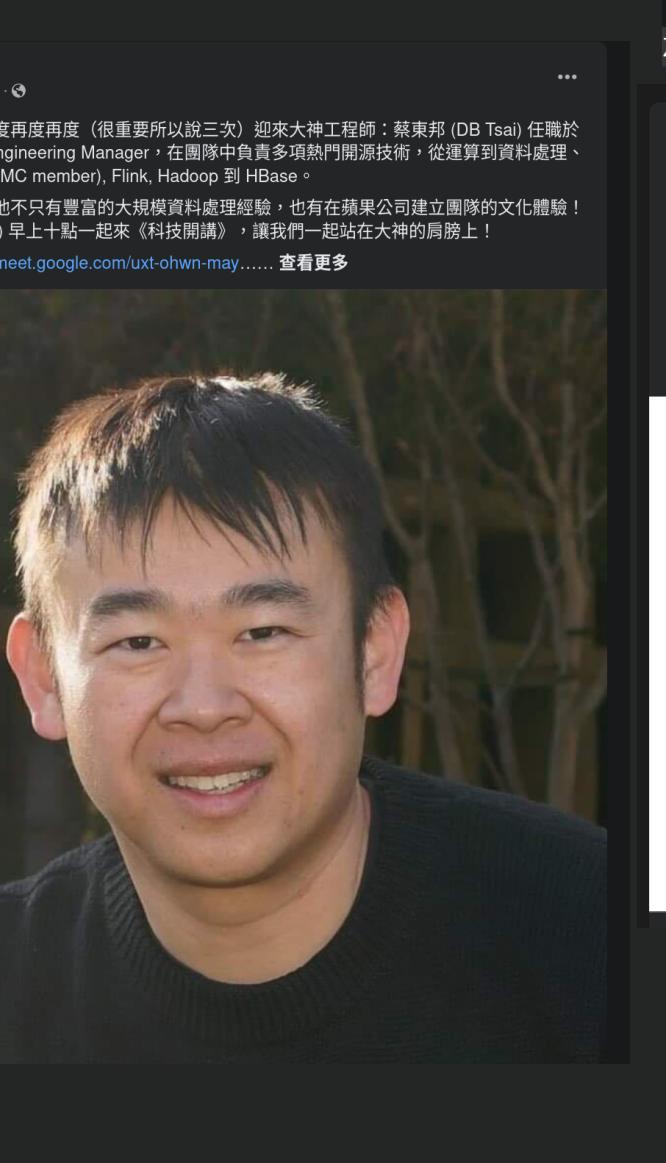
Creator / Maintainer @ Magisk
GitHub stars: 46.8K
Downloads: 270M

BS @ NTUEE; MS @ UMich

Senior SWE @ Google Android

John Wu
topjohnwu
Unfollow
At 9.5k followers · 0 following

Google Meet:
<https://meet.google.com/uxt-ohwn-may>



源來適你

6月13日

本週六「科技開講」延續職場時間！上周六我們看了高大上的臺灣產業報告，這週我們聚焦個人來討論討論如何「面試」！

本周的講者：王超，現任的 Meta 工程師，在社招和校招面試中共拿過 17 個 offers (Google, Meta, Amazon, Tiktok, hot startups ... etc)，做過 10+ 場 coding 面試的面試官。喜歡爬山，唱卡拉OK 和分享心理學的科普講座！這次講者不藏私分享以下幾個主題：

- 美國軟體工程師面試流程
- 準備面試的 best practice..... 查看更多

科技開講：美國求職教戰手冊 - 王超

- 現 Meta 工程師
- 社招和校招中共拿過 17 個 offers (Google, Meta, Amazon, Tiktok, hot startups ... etc)
- 大綱：
 - 美國軟體工程師面試流程
 - 準備面試的 best practice
 - 投遞時間和關鍵步驟的建議

獲取工作 - no 刷題

podcast.opensource4you.tw



8:40



9:00



1:55



Luke (陳孝勇) Chen · 1 度
Software Engineer



Tzu-Li (Gordon) Tai ✅ · 1 度
Senior Engineering Manager at Confluent, Apache Flink Committer and PMC Member



union
The AI Orchestra
(Kevin) Huan-Ping Su ✅ · 1 度
Apache Submarine PMC at The Apache Software Foundation



Guozhang Wang - 2020年10月20日 星期二 凌晨1:24:34 [GMT+8]

Hello all,

I'm happy to announce that Chia-Ping Tsai has accepted his invitation to become an Apache Kafka committer.

Chia-Ping has been contributing to Kafka since March 2018 and has made 74 commits:

<https://github.com/apache/kafka/commits?author=chia7712>

He's also authored several major improvements, participated in the KIP discussion and PR reviews as well. His major feature development includes:

- * KAFKA-9654: Epoch based ReplicaAlterLogDirsThread creation.
- * KAFKA-8334: Spiky offsetCommit latency due to lock contention.
- * KIP-331: Add default implementation to close() and configure() for serde
- * KIP-367: Introduce close(Duration) to Producer and AdminClients
- * KIP-338: Support to exclude the internal topics in kafka-topics.sh command

In addition, Chia-Ping has demonstrated his great diligence fixing test failures, his impressive engineering attitude and taste in fixing tricky bugs while keeping simple designs.

Please join me to congratulate Chia-Ping for all the contributions!

-- Guozhang



Matthias J. Sax - 2020年10月20日 星期二 凌晨1:26:29 [GMT+8]

Congrats Chia-Ping!



Bill Bejeck - 2020年10月20日 星期二 凌晨1:31:55 [GMT+8]

Congratulations Chia-Ping!

-Bill



Sophie Blee-Goldman - 2020年10月20日 星期二 凌晨1:39:14 [GMT+8]

Congrats!



Bruno Cadonna - 2020年10月20日 星期二 凌晨1:40:34 [GMT+8]

Congrats!

Best,

Bruno



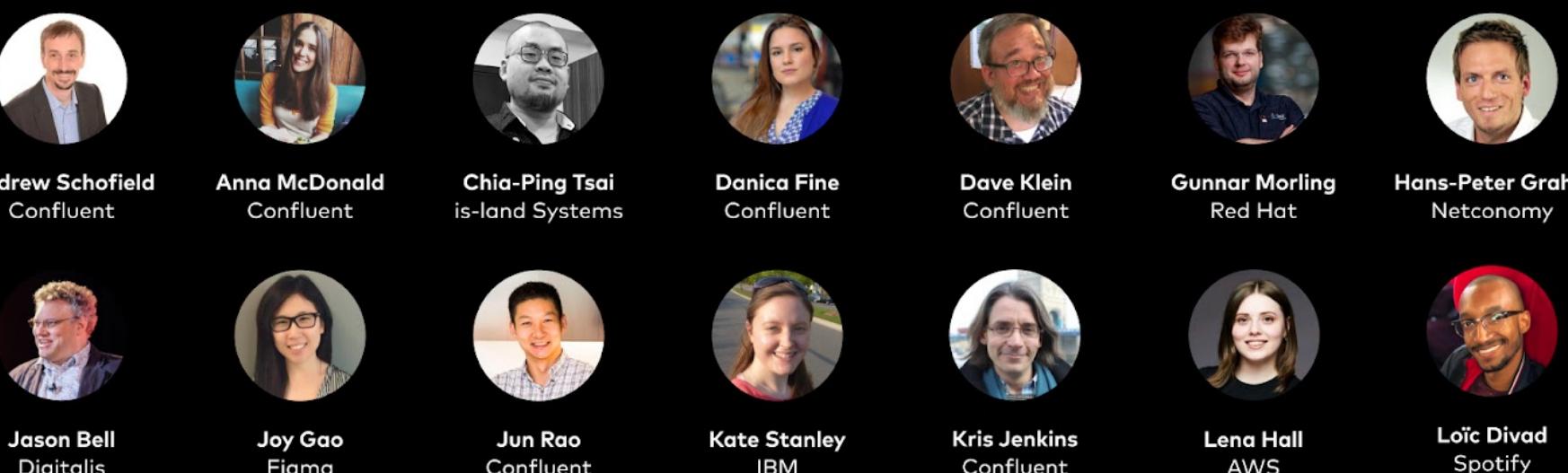
Kowshik Prakasam - 2020年10月20日 星期二 凌晨1:51:30 [GMT+8]

Congrats!

Cheers,
Kowshik

chhsiao	Chun-Hung Hsiao	mesos-pmc , mesos
chia7712	Chia-Ping Tsai	hbase-pmc , hbase , incubator , kafka-pmc , kafka , member , yunikorn-pmc , yunikorn
chico	Chico Charlesworth	incubator

KAFKA SUMMIT PROGRAM COMMITTEE 2022



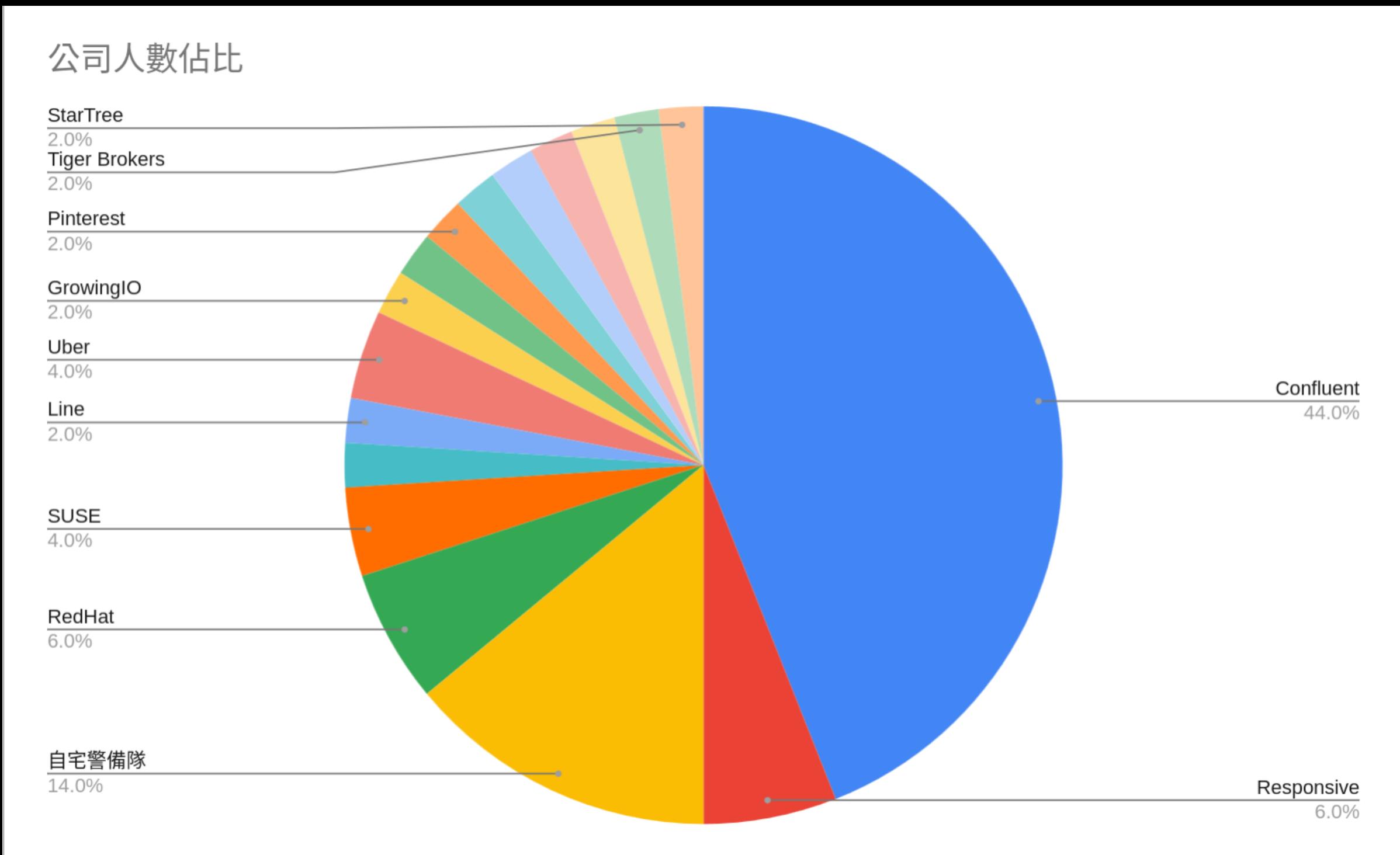
實現自我 - 取得國際專案頭銜



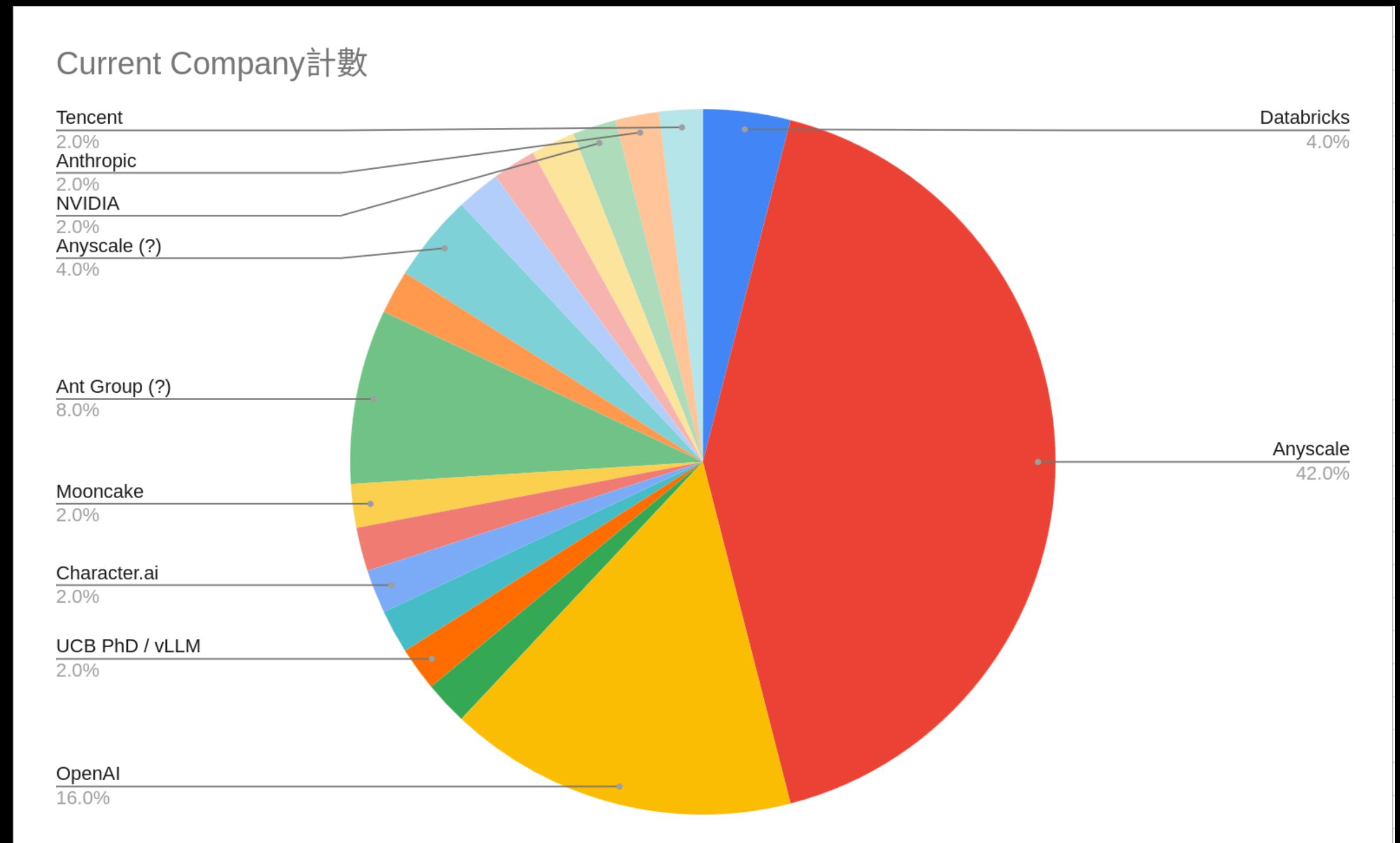
為何公司喜歡參與開源？

top 50 貢獻者的公司分布

Apache Kafka



Ray



The Mark Zuckerberg Interview



- ① **吸引頂尖人才：**許多優秀的軟體工程師和研究人員都熱衷於參與開源專案。Meta 積極擁抱開源，特別是在 AI 領域開源其模型如 Llama，這讓它在這些人才眼中更具吸引力。對開發者來說，為一個開放的、有影響力的專案貢獻代碼，或是使用開源工具，是他們職業發展的重要部分。
- ② **激勵內部員工：**允許或鼓勵員工為開源專案做出貢獻，可以提高他們的工作滿意度和歸屬感。當他們的成果能夠被廣泛使用和改進時，會更有成就感。
- ③ **擴大潛在人才庫：**由於 Meta 的 AI 模型是開源的，全球有更多的開發者能夠學習、使用和改進這些技術。這實際上是在培養一個龐大且熟悉 Meta 技術的潛在人才庫，未來 Meta 在招聘時，可以從中找到更多具備相關經驗的候選人。
- ④ **促進知識交流：**開源社群是一個巨大的知識分享平台。Meta 不僅能從外部社群獲得貢獻，也能通過與社群的互動，提升自身團隊的技術水平和創新能力。

Josep Prat – Sustainability beyond funds: Extrospective OSPOs

Senior Engineering Director



[Https://youtu.be/gYIzeL2WxQA?si=xWDU6UjZYasaiZY7](https://youtu.be/gYIzeL2WxQA?si=xWDU6UjZYasaiZY7)

要解決的問題：

- 許多專案依賴大量的開源函式庫，這些依賴關係複雜且難以管理，可能導致潛在的風險和維護問題

解決方式：

- 企業成立一個專門的工程師團隊（隸屬於開源專案辦公室），擁有自己的時間，工作完全專注於開源專案。選擇對企業業務影響最大的關鍵依賴專案，優先投入資源維護

對公司的好處：

- 確保企業依賴的專案能夠長期存在和發展
- 更快的應對安全風險：例如 Log4j 事件
- 引導專案朝公司所需的方向發展
- 增加專案的貢獻者，避免過度依賴單一維護者
- 鼓勵更多企業參與貢獻，形成更健康的開源社區

實際案例

- Aiven 成立了一個約 15 人的團隊，專注於他們提供的雲端服務所依賴的六個主要開源專案（Kafka, Flink, Cassandra, PostgreSQL, OpenSearch, ClickHouse），並培養了多位專案維護者

商業公司競爭「規格」的制定權

Confluent v.s Aiven v.s AutoMQ v.s Salesforce



源來適你

5月14日上午2:17 · ●

...

大家心裡或許都曾有過一個疑問：《源來適你》的 Kafka 社群為何如此熱絡？其實原因很簡單：作為開源技術的領航社群，我們不僅希望大家能透過參與開源技術獲得學習成長，更期盼這些技能能夠直接與就業市場接軌。因此，廣泛應用且需求殷切的 Kafka 自然成為《源來適你》的主力項目之一。

正如 AI 所稱譽：「Apache Kafka 是首屈一指的即時數據串流平台。根據多份報告顯示，其市場佔有率約在 26.7% 至 38.49% 之間。包括財富 100 強企業中約 80% 的公司，都已在不同程度上部署了 Kafka。」在如此高的市場熱度下，提供 Kafka 商業服務和就業機會的公司自然如雨後春筍般湧現，例如 Confluent、紅帽、Responsive、AutoMQ、Aiven 和 RedPanda 等等。

然而，商業公司的蓬勃發展是否全然是好事？倒也未必。一個最典型的例子便是，在重要的技術決策上，不同的聲音難免會讓討論變得更加複雜。以近期備受關注的 Kafka 存算分離議題為例，截至目前為止，已有三家公司提出了各自的解決方案：

- * Aiven: KIP-1150 Diskless topics
- * Slack: KIP-1176 Tiered storage for active log segment
- * AutoMQ: KIP-1183 Unified shared storage

此外，包括 Confluent 和 WrapStream 等公司，據悉也已具備相關方案，只是尚未正式表態。

沒有人樂見社群發展受阻，但同時，也沒有人希望自己的方案不被採納。推動社群方案成為業界標準，始終是一個極具吸引力的目標。那麼，這場關於 Kafka 存算分離的競逐，Kafka 的下一個黃金十年，最終將會如何落幕呢？小編會持續為大家追蹤報導！

開源的商業模式

有三種主要的服務模式在開源軟體領域中被廣泛採用：

- ① 100% 開源 + 訂閱服務：這種模式的核心是將所有技術完全開源，並透過提供訂閱、服務和培訓來獲取收益。
- ② Open Core 模式：此模式僅將技術框架開源，而將關鍵部分閉源並作為增值內容以實現商業價值。
- ③ SaaS 模式：將開源軟體部署到公有雲上，以雲服務為用戶提供服務，進而實現商業變現。

100% 開源 + 訂閱服務

對於企業用戶而言，「社群版軟體」的更新速度過快，若要採用，則需投入較多的資金和精力進行維護，特別是安全性問題更是令人擔憂。由於開源社群無法協助企業用戶解決上述問題

紅帽將「社群版軟體」增加了企業級應用與服務，例如提供全年無休（7x24小時）的電話服務，以及過往問題的解決方案等，然後以訂閱模式提供給客戶。

Open Core 模式

- ① Confluent 為 Apache Kafka 保留為商業專有軟體。需要使用相關功能的客戶，要麼自行開發實現或尋找其他開源實現，要麼只能向 Confluent 購買相關的商業服務
- ② DataStax 基於開源的 Apache Cassandra 和 Apache Pulsar 開發了專有產品。
- ③ Starburst 基於開源的 Presto/Trino 開發了專有產品。
- ④ EnterpriseDB 提供與 PostgreSQL 相關的商業專有軟體。

SaaS 模式

- ① Aiven 提供 Apache Kafka 的雲端服務。他們聘用開源軟體的開發者，負責回應內部使用時遇到的問題，並將錯誤修復和功能改進推送到上游專案
- ② Tetrate 基於開源的 Istio 和 Envoy 開發了專有產品，並提供了商用的 Istio 發行版。
- ③ Firebolt 基於開源的 ClickHouse 等軟體開發了專有產品。
- ④ Decodable 基於開源的 Apache Kafka 和 Apache Flink 開發了專有產品。

我能如何參與 開源？

任何方式都可以 推廣、參加活動、發 patch、下載來用

參與國際開源專案的四個重點

當個使用者，下載來用

APACHE QUICKSTART

Everything you need to know about Kafka in 10 minutes (clicking the image will load a video from YouTube)

What is Apache Kafka?

STEP 1: GET KAFKA

Download the latest Kafka release and extract it:

```
1 | $ tar -zxf kafka_2.13-3.7.0.tgz
2 | $ cd kafka_2.13-3.7.0
```

回報你發現的問題

Kafka / KAFKA-16604

Deprecate ConfigDef.ConfigKey constructor from public APIs

Type: Improvement Status: OPEN

Priority: Major Resolution: Unresolved

Affects Version/s: None Fix Version/s: None

Component/s: None Labels: None

Description

Currently, one can create ConfigKey by either invoking the public constructor directly and passing it to a ConfigDef object or by invoking the a bunch of define methods. The 2 ways can get confusing at times. Moreover, it could lead to errors as was noticed in KAFKA-16592. We should ideally have only 1 way exposed to the users which IMO should be to create the objects only through the exposed define methods. This ticket is about marking the public constructor of ConfigKey as Deprecated first and then making it private eventually.

Attachments

Drop files to attach, or browse.

Comments

Activity

Comments

Chia-Ping Tsai added a comment - 23/Apr/24 20:00

Sagar Rao Thanks for opening this jira. Do you have free cycle to take over it?

Slack

In order to see discussions, first confirm access to your Slack account(s) in the following workspace(s): ASF

貢獻程式碼

KAFKA-16310 ListOffsets doesn't report the offset with maxTimestamp a... #15621

Merged chia7712 merged 52 commits into apache:trunk from chia7712:KAFKA-16310 2 weeks ago

Conversation 87 Commits 52 Checks 0 Files changed 19

chia7712 commented last month

We do iterate the records to find the `offsetOfMaxTimestamp` instead of returning the cached one when handling `ListOffsetsRequest.MAX_TIMESTAMP`, since it is hard to align all paths to get correct `offsetOfMaxTimestamp`. The known paths are shown below.

1. `convertAndAssignOffsetNonCompressed` -> we CAN get correct `offsetOfMaxTimestamp` when validating all records
2. `assignOffsetsNonCompressed` -> ditto
3. `validateMessagesAndAssignOffsetsCompressed` -> ditto
4. `validateMessagesAndAssignOffsetsNonCompressed` -> ditto
5. `appendDataForTopicAppendAndValidateRecords` -> we CAN'T get correct `offsetOfMaxTimestamp` as iterating all records is expensive when fetching records from leader
6. `LogSegmentRecover` -> ditto

<https://issues.apache.org/jira/browse/KAFKA-16310>

Committer Checklist (excluded from commit message)

Verify design and implementation

不斷貢獻直到取得開源頭銜

Apache / kafka committer Available projects (372):

Archives (2011 - 2024):

[ANNOUNCE] New Kafka PMC Member: Bruno Cadonna

Guozhang Wang Hi everyone, I'd like to introduce our new Kafka PMC member: Bruno. Bruno has been a committer since April, 2021 and has been very active in the community. He's a key co...

[ANNOUNCE] New Kafka PMC Member: Bruno Cadonna

Guozhang Wang 2023-11-01 14:40 +08:00

I'd like to introduce our new Kafka PMC member, Bruno. Bruno has been a committer since April, 2021 and has been very active in the community. He's a key contributor to Kafka Streams, and also helped review a lot of horizontal improvements such as Mockito. It is my pleasure to announce that Bruno has agreed to join the Kafka PMC.

Congratulations, Bruno!

-- Guozhang Wang, on behalf of Apache Kafka PMC

Bill Bejeck 2023-11-01 14:40 +08:00

Congrats Bruno! Well deserved.

Chris Egerton 2023-11-01 14:40 +08:00

Congrats!

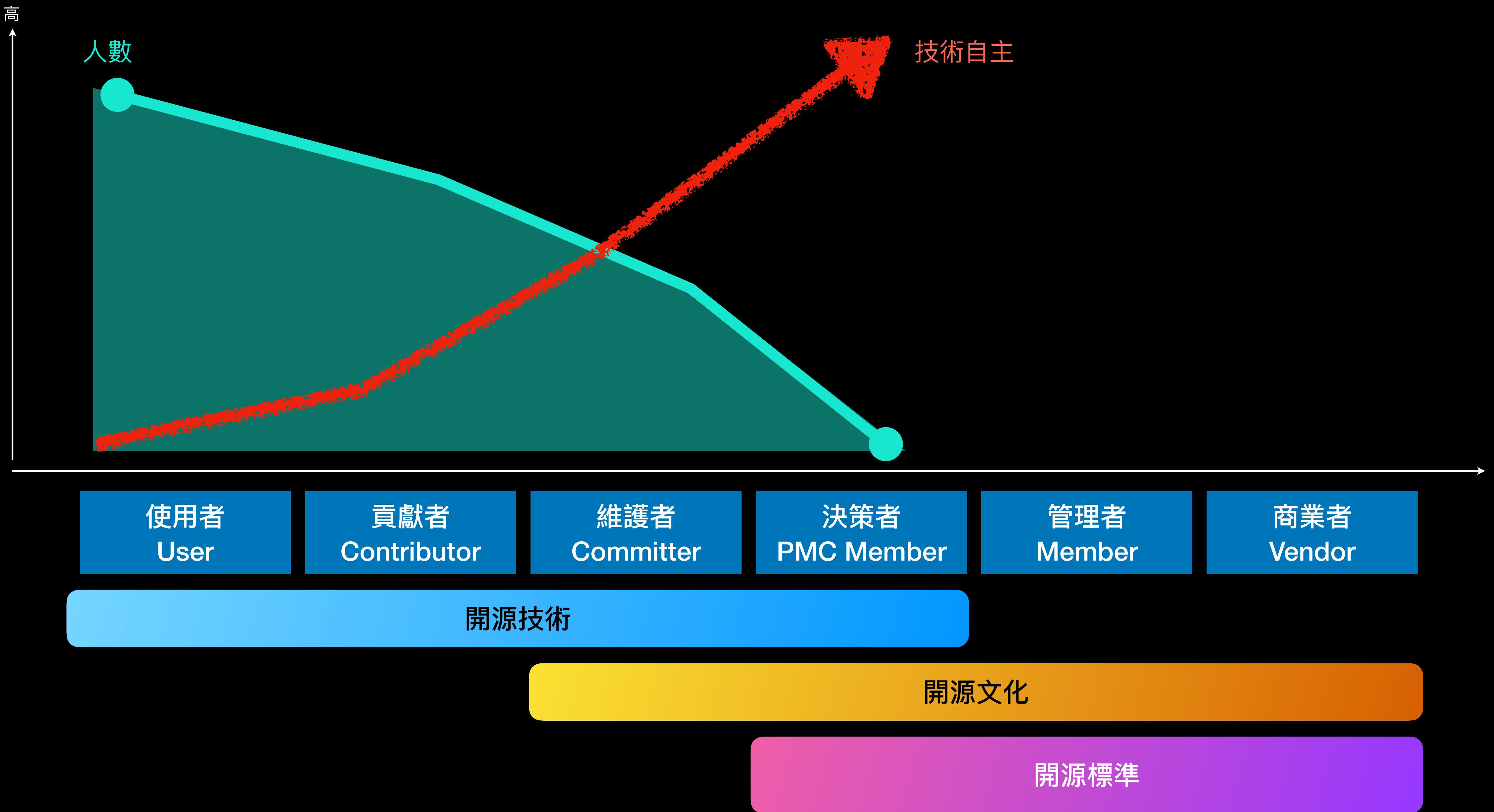
Lucas Brutschy 2023-11-01 14:40 +08:00

Wow, congratulations!

John Roessler 2023-11-01 14:40 +08:00

Congratulations, Bruno!!!

台灣離開源的距離有多遠？



開源相關的數據比較

	美國	中國	台灣
國際 開源基金會	Apache Linux CNCF	開放原子開源 基金會	0
Apache 專案數量	一大堆	>10	0
Apache 董事	一大堆	3	0
Apache Member	一大堆	>50	<10
Apache 分部	本部	>5	0 ... 1(?)
開源風氣	自然而然	國家支持	?



你想要取得熱門技術的頭銜？

要跟一群大公司的資深員工競爭
你的臉沒存在感沒人會理你
你很難累積有價值的貢獻
要投入非常多的時間
日夜作息顛倒
沒有標準

I also reviewed many of his pull requests and I was never really satisfied by them. I found that he often missed fundamental things while working on the new consumer.

I find him a little sloppy in PR reviews. There is frequent occasions, in which he approves a PR, and I review it afterward, and still leave a lot of comments, and we need one or two more rounds with the contributor before we can actually merge the PR.

熱門技術的頭銜很不容易

Working with him on many of his PRs I sometimes feel like he did not fully think through the changes he's making, or he missed a small part of the task.

I think that he does not meet the bar yet from a technical point of view. I don't trust him to merge code at this point given the quality of his reviews.

I do not believe xxx meets the technical bar for committership. His contributions related to xxxx have been seriously underwhelming. I'd be afraid for the quality of our code base if he was allowed to merge changes.

我能夠如何建立開源？

Community and Foundation

世界三大開源基金會

	Apache 基金會 (ASF)	Linux 基金會 (LF)	CNCF 基金會 (CNCF)
成立時間	1999 年	2000 年	2015 年
宗旨與使命	促進開源軟體的發展和維護，強調「社群優於程式碼」(Community Over Code) 的理念，提供法律保護和品牌管理。	促進 Linux 內核及其他開源技術的發展，為全球開源協作提供中立的平台和基礎設施。	推動雲原生計算技術的普及和發展，專注於構建一個可持續的雲原生生態系統。
典型項目	Apache HTTP Server, Hadoop, Kafka, Cassandra, Spark, Flink, Tomcat, Airflow, Zookeeper 等。	Linux Kernel, Node.js, Hyperledger, OpenSSF, PyTorch, RISC-V 等。	Kubernetes, Prometheus, Envoy, etcd, containerd, Argo, Cilium 等。
授權協議	主要使用 Apache License 2.0，這是一個寬鬆的開源許可證，允許商業使用和修改。	旗下項目可能使用多種開源許可證，包括 GPL、MIT、Apache License 等，視具體項目而定。	旗下項目主要使用 Apache License 2.0，也可能有其他兼容開源許可證。

自己經營 v.s 捐給基金會

	自己營運開源專案	捐給基金會營運
控制權	完全控制	共享控制，需遵循基金會規則
決策效率	高	可能較慢，需達成共識
營運負擔	重，需處理多方面事務	輕，基金會處理非開發事務
資金與資源	資源有限，需自行籌措	資源更豐富，可獲得更多贊助和支援
社群發展	挑戰大	更容易吸引和擴大社群
可信度	依專案自身積累	基金會背書，顯著提升
長期可持續性	依賴核心團隊，風險較高	有基金會支持，更具備長期可持續性

Apache Kafka 發展時間圖

- 2010 年：LinkedIn 內部開發
Kafka 最初由 LinkedIn 的工程師團隊 (Jay Kreps, Neha Narkhede, Jun Rao 等人) 在內部開發，旨在解決 LinkedIn 大規模數據流的即時處理和低延遲需求。它被設計為一個分佈式日誌系統，用於收集和傳遞高吞吐量的日誌數據。
- 2011 年初：LinkedIn 將 Kafka 開源
LinkedIn 決定將 Kafka 開源，並於 2011 年 1 月左右在 GitHub 上公開。這是一個重要的里程碑，標誌著 Kafka 從內部工具走向更廣闊的開源社群。
- 2011 年 7 月 12 日：進入 Apache 孵化器 (Incubator)
開源僅僅數月後，Kafka 被 Apache 軟體基金會 (Apache Software Foundation, ASF) 接受，正式進入其孵化器專案。這意味著 Kafka 將在 Apache 的治理模式下成長，目標是建立一個強大而多元的社群，並遵循「Apache Way」。
- 2012 年 10 月 23 日：從 Apache 孵化器畢業，成為頂級專案 (Top-Level Project, TLP)
經過約 15 個月的孵化期，Kafka 成功畢業，正式成為 Apache 軟體基金會的頂級專案。這標誌著 Kafka 已經證明其擁有一個活躍的社群、穩定的開發流程和健全的治理模式，達到了 Apache 基金會對頂級專案的標準。
- 2013 年：Kafka 0.8 版本發佈
在成為 Apache TLP 之後不久，Kafka 0.8 版本發佈，這是它作為 Apache 頂級專案後的重要版本，包含了內置的數據複製等關鍵功能。
- 2017 年 11 月 28 日：Apache Kafka 1.0 版本發佈
雖然 Kafka 在 2012 年就成為了頂級專案，但直到 2017 年才發佈了 1.0 版本，這標誌著其作為一個成熟的事件串流平台的里程碑。此後，Kafka 的發展速度更快，並被業界廣泛採用。

捐專案給 Apache 基金會

- ① 了解 Apache Way 和 ASF 文化：Apache 基金會強調「社群勝於程式碼」（Community Over Code）的理念。這意味著一個健康、開放、多元且共識決的社群對於專案的成功至關重要。
應問題、解釋疑慮並根據反饋修改提案。這是一個開放透明的過程，所有的討論都會在郵件列表中公開進行。
- ② 尋找導師（Mentors）：這是至關重要的一步。您需要找到至少三位活躍的 Apache 基金會成員，他們願意擔任您專案的導師。
⑤ 孵化階段（Incubation Phase）：如果提案獲得孵化器 PMC 的投票通過，您的專案就會進入孵化階段，您的專案社群將在導師的指導下，逐步適應 Apache 的治理模式、基礎設施和運營方式。
- ③ 準備專案提案（Project Proposal）：為什麼這個專案應該加入 Apache 基金會？它如何符合 Apache 的目標和價值觀？初步目標：
⑥ 畢業（Graduation）：當專案社群足夠成熟，並且符合 Apache 的所有要求時，就可以申請「畢業」。畢業意味著專案成為 Apache 基金會的一個頂級專案（Top-Level Project, TLP）
- ④ 提交提案並進行討論：提案會引起孵化器 PMC 和其他社群成員的討論。您需要積極回

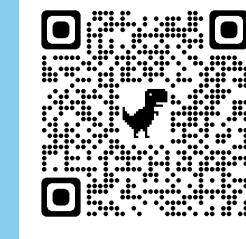
什麼是《源來適你》？



源來適你

0 入場費

0 門檻 只缺想展現影響力的你



10+ 工程師榮獲國際專案頭銜

2024 Apache 國際研討會
唯一台灣民間組織

熱門開源技術 Apache Kafka, Ozone,
KubeRay ... 第二大貢獻組織

10+ 每週線上開源技術會議

參與人數
500+
線下交流活動

開源貢獻
2000+
Commits

100+ 國內外專家技術分享



如何透過開源社群「源來適你」找到工作

大家好，我是羅昱翔，目前在CMU攻讀Data Science。這篇文章想分享我如何利用開源社群找到實習的經驗。從去年8月到今年1月，我投了389份職職。上半年幾乎沒有拿到面試機會，少數面試也憑實力被拒。直到後來加入了「源來適你」社群，由楷訓帶領的KubeRay組，開始接觸開源，才讓我的履歷、經歷與技術能力提升。更重要的是，我因此建立了一些關鍵的connection，開始接觸更多潛在的工作機會。最終，我成功憑藉開源經驗拿到了CrowdStrike MLE Intern和Samsara SWE Intern的Offer。

首頁 » 文章

貢獻開源並不難：我透過「源來適你」社群成為 Flyte 核心貢獻者！

這篇文章講述了我這幾個月貢獻 Flyte 這個開源專案，到最後成為核心貢獻者的經歷分享。想要貢獻開源但是卻怕自己實力不夠？加入台灣開源社群「源來適你」，藉由社群內導師的帮助可以讓你克服恐懼、循序漸進的成為開源專案核心貢獻者。

發佈於: 2024-08-18 · 最後修改: 2024-09-03 · 劉奇聖

致謝

感謝Kevin，週會都給很多建議，也會跟我約1-1討論PR

感謝Han-Ju，引導我進入開源的世界，還提名我當committer

感謝一起做開源的大佬們(Chi-Sheng, Austin, Chun-Mao, ...)，週會都聽到很多酷東西，他們做的feature都很潮

感謝嘉平host源來適你這個開源社群，每週邀請講者來分享

源來適你

源來適你是一個在台灣實際貢獻開源的非營利組織

裡面有許多 Mentor 帶你實際貢獻開源專案，包括 Apache Airflow, Apache Kafka, Apache YuniKorn 等等

這邊都是以中文溝通，如果有問題可以更自在的用中文提問

• 「源來適你」目前有 Mentor 帶的專案列表

• 關於源來適你更深入的介紹 Kafka Community Spotlight: TAIWAN 🇹🇼 by Stanislav's Big Data Stream

因為這件事讓我重新思考我的選擇，即使我進了夢想公司，但如果有一天被火，那我的職涯還剩下什麼？有什麼 Title 是公司拿不掉的？剛好這時候看到了嘉平學長在舉辦「源來適你」的活動，當初在學校時就有聽朋友說他們實驗室的學長很厲害，是Apache Hadoop PMC，所以就上 LinkedIn 肉搜了一下，然後才在學長的履歷中發現 YuniKorn 這個專案。YuniKorn 可以作為 Kubernetes Scheduler，非常適合 Big Data & ML 的 workload，由於我在工作中做過 Kubernetes 的開發，對 Golang 也算熟悉，所以對我來說 YuniKorn 是很好的起點，我第一個貢獻是 2023/05/09 提交的，後來到了 7 月初，不喜歡新的公司，所以就索性辭職 fulltime 做貢獻，最終在 2023/08/21 時收到邀請成為 YuniKorn Committer 的信，期間貢獻了四十幾個 PR。

不到兩年培育 N+ 開源人才

參與者心得

- [The Apache YuniKorn committer's path](#)
- [2024 Apache CommunityOverCode 亞洲場遊記](#)
- [貢獻開源並不難：我透過「源來適你」社群成為 Flyte 核心貢獻者！ | 劉奇聖](#)
- [開源菜雞入門心得](#)
- [我的 Open Source 之路](#)
- [參與 YuniKorn 貢獻的心得](#)
- [Open Source 開源社群的第一門課 | 如何成為 Apache Committer](#)
- [Flyte 開源之旅 — 從開始到現在](#)
- [如何成為 Open Source Committer \(Flyte\) | 大學生 21 周心得分享](#)
- [普通人也能做到的 Open Source 開源一年 100+ PRs 完整指南 feat. 源來適你, Flyte](#)
- [開源入門到加入夢幻新創！](#)
- [不尋常的找工作方式：我透過貢獻 KubeRay 開源專案找到美國跨國遠端軟體工程師職缺！ | 劉奇聖](#)
- [如何透過開源社群「源來適你」找到 Samsara 工作](#)
- [從歷史系到阿帕契 - 轉職仔的開源之路](#)
- [從 0 成為 Apache Airflow Committer](#)
- [Open Source 入門心得](#)
- [Apache O](#)

我很幸運，在焦慮的海浪裡看到了岸，於是我在軟體工程裡不斷革新的浪潮裡被淹沒，也找到了努力的方向，如今也達成了一個小小的里程碑。所以下來我會寫一些關於這一年沒寫部落格之後我在做什麼，我是如何透過開源社群「源來適你」投入到開源貢獻，進而取得開源技術頭銜。

加入開源的契機

第一次了解到開源是在 Terry 的影片，裡面提到了源來適你這個社群。這個社群的好處是如果對 issue 或 PR 有任何的問題，都可以在群組內提問，大家都會很熱心的幫忙解決問題，不會出現孤立無援的情況。當時了解到這個社群時就馬上加入了。

直到蔡嘉平學長籌組了“源來適你”，並要在2023年10月時舉辦“香菜饅頭營”，這個活動會由 Flyte, Ray, YuniKorn, Ozone 等開源專案的技術專家帶領，用4~5個月的時間幫助有心貢獻的開發者參與開源。同時，我也正面臨一些職涯瓶頸，感受到冒牌者症候群的困擾，總覺得“別人都好強，我怎麼還是這麼廢”，一直在思考要如何突破。於是最後決定要抓住這個機會，作為一個我想跳出舒適圈的藉口。在評估當下工作的負擔後，覺得我可能沒辦法同時參與饅頭營活動又專心把工作做好。最後就決定提離職，把這次的活動當作我的一個Career break。

大約是從大學時期開始接觸開源吧，還記得當時看到Firefox網路貼文便因為好奇參加Firefox的小型聚會，並貢獻中文官方文件，當時只覺得很酷，對於開源這件事對我有多大幫助反而沒有太多想法。

畢業後，繼續參與幾個開源專案：HBase、Trino等，從這階段理解到即便只有“追”程式碼，也能提昇個人技術水平，但也因為見識到更多事物後，了解到自己究竟還離目標差距多遠，也深刻理解到人外有人、天外有天。

直到蔡嘉平籌組了源來適你，我在2023年9月也跟著加入貢獻YuniKorn的行列，當時只覺得貢獻YuniKorn除了可以熟悉K8s，同時又可以學會Golang，秉持著做中學的想法，便決定馬上投入。想當然，剛開始便面臨了挑戰，大概快一個月熟悉YuniKorn各個元件的運作機制。但理解專案是基本，找議題才是最關鍵的...好在根據自己以前貢獻其他的開源專案的經驗，以及諮詢各路好手的建議，最終歸納出找議題的方法大概有以下幾項：

直到有一天，在嘉平的臉書，看到『源來適你』的開源活動，就報了名去參加會議，但可能過去的陰影還是在，始終沒有填寫饅頭營問卷（還有部分是不少新的開源技術，連聽都沒聽過，感覺自己好像老了XD）

然後又過了4個月，忽然有一天聽到源來適你Podcast，Gordon講的一句話：

“不要因為跟那個人覺得差距這麼大，你就連第一步都不願踏出去”

就覺得該勇於去嘗試，於是我就主動去敲翰儒，自我介紹一番，然後翰儒又把我邀請至Flyte台灣群組中，我的Flyte旅程起點就此展開。

謝謝嘉平創辦源來適你，總是即時的回答我遇到的所有問題，你示範了“真正的體貼是讓人察覺不到的”這句話，希望我們之後也能當朋友

謝謝桓平，這一年來總是用行動照顧我，從你身上學到很多成熟的想法和特質，(大多時候我是用感受的，那是由許多想法行為形成的一個氣質)，以及看到了什麼是熱情，以後跟 Terry 還有 TheKellyYang 一起拍片時一定揪你



源來適你

1月5日

2023年叮叮噹饅頭順利落幕！感謝眾多講者為我們帶來 YuniKorn, Flyte, feature-toggle, Ray, and Accio 等精彩的技術分享！同時也要非常非常感謝 Dcard Tech Studio 強大的場地支援！

大家 & 開源，我們 2024 年再相會！！！



[查看洞察報告和廣告](#)

[加強推廣貼文](#)



源來適你

2023年11月21日

小饅頭營從南到北的校園巡迴，參與學生超過三百位 (誤差範圍3%)，非常感謝眾多開源領域的前輩投入時間，以下就精選前輩們笑得燦爛的照片作為小饅頭營的謝幕 😊



參與人數
500+
線下交流活動



國際 Apache 研討會

唯一台灣民間組織



COMMUNITY
THE ASF CONFERENCE
CODE

CommunityOverCode Asia 2024

⌚ July 26-28, 2024 ⌚ Hangzhou, China

State of Apache YuniKorn

 Yu-Lin Chen Apache YuniKorn committer	 Chia-Ping Tsai Apache Member, Apache HBase PMC member, Apache Kafka PMC member, Apache YuniKorn PMC member
 Kuan Po Tseng Apache YuniKorn committer	 Ting Yao Huang Apache YuniKorn PMC member
 PoAn Yang Apache YuniKorn committer	

Scan the QR code or visit asia.communityovercode.org to join us





國外科技媒體採訪

許多開源知名大人物轉發

BIG DATA STREAM

kafka community spotlight: TAIWAN 🇹🇼

6 high-impact Kafka contributors from Taiwan share their exprience

FEB 21, 2025



Share

Kafka Community Spotlight

TAIWAN



Mickael Maison @MickaelMaison · 1h

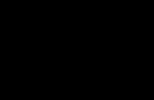


...

Not all heroes wear capes.

It's hard to truly describe the impact of Chia-Ping on Apache Kafka.

One example: Chia-Ping is not only consistently the top reviewer but if we look at the last quarter he has made as many code reviews as the next 8 top reviewers combined!



Stanislav Kozlovski ✅ @BdKozlovski · Feb 21

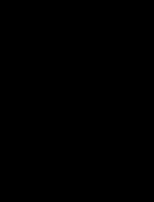
This is the most impactful Apache Kafka mentor you've never heard of:

Chia-Ping Tsai.

...

Show more

Show this thread



1



3



97



Stanislav Kozlovski ✅ · 1 個月前
The Kafka Guy
2 個月前

This is the most impactful Apache Kafka mentor you've never heard of:

Chia-Ping Tsai.

In just 18 months, he bootstrapped a large Taiwanese open source community boasting:

- 5000 participants
- 15,000 Slack messages/month
- 10 meetings/week
- 20+ Apache committers

Together, this community has merged more than 591 PRs in Apache Kafka! 🇹🇼

He did all of this while being a committer and PMC member in Apache Kafka, Apache HBase and Apache Yunikorn 🤓

Just in Kafka he's created around 600 issues and reviewed more than 1,600 commits 😊

And his next task?

Help his closest community members become Apache Kafka committers. 👍

I accidentally stumbled upon this community while reviewing Kafka PRs. I saw a lot of very active contributors from Taiwan and one common link - Chia-Ping.

I joined their Slack and was blown away by the activity. 😊

I then conducted a group interview with [Chia-Ping Tsai](#) and 5 other very active members from the Kafka community. ([TengYao Chi](#), [Po-An Yang](#), [Jiunn-Yang Huang](#), [Kuan-Po \(Cooper\) Tseng](#) and [Luke Chen](#)) 🙌

We spoke about all sorts of things, including biggest pet peeves in Kafka, hardest challenges each one faced, lessons from contributing to open source and much more!

Check out the full interview here:

✓ https://lnkd.in/eH_xwFma

If you'd like to see the Kafka community keep growing like this - boost this post by reposting to your network! 💥

#Kafka #ApacheKafka #OpenSource #OpenSource4You

顯示翻譯



200 您和其他 476 人

25 則回應 · 21 則轉發

台灣空前絕後的 Apache 分部

獲得超過 20 票 (同期另一個國家只有5票)

[ALC] Request to establish ALC in Taipei

Posted to dev@community.apache.org



Chia-Ping Tsai - 2025年4月3日 星期四 晚上11:21:55 [GMT+8]

hi All,

I'm an Apache Member (Apache ID: chia7712), and I've led the local open-source community, 源來適你/opensource4you [1], for two years. We've successfully fostered contributions to various Apache projects, including Apache Kafka, Apache YuniKorn, Apache Airflow, and Apache Gravitino, resulting in over 10 community members becoming Apache committers.

Leveraging the Apache Local Community (ALC) name would be a valuable asset in our efforts to attract more contributors to open-source software projects.

Thanks,
Chia-Ping

[1] <https://www.facebook.com/opensource4you>

The grid contains 20 comments from various Apache committers:

- Zhe You Liu - 2025年5月20日 星期二 下午5:46:44 [GMT+8]: I'm an Apache Airflow Committer. I hope we can have a formal Apache community in Taipei. Thanks, Zhe-You Liu
- TengYao Chi - 2025年4月10日 星期四 清晨6:26:17 [GMT+8]: I'm an Apache Kafka Committer. Thanks Chia-Ping for running this incredible community! Can't wait to see ALC Taipei coming true!
- Jax Liu - 2025年4月9日 星期三 凌晨12:58:57 [GMT+8]: I'm an Apache DataFusion Committer. Thanks Chia-Ping. It sounds great. Jax
- Wei Lee - 2025年4月4日 星期五 下午4:39:34 [GMT+8]: I'm a committer of Apache Airflow. Would love to see such a community happen in Taiwan. Best, Wei
- Chi-Sheng Liu - 2025年5月20日 星期二 下午4:58:54 [GMT+8]: I'm a Committer of Apache Submarine. Looking forward to having Taipei ALC! To unsubscribe, e-mail: dev-unsubscribe@community.apache.org For additional commands, e-mail: dev-help@community.apache.org
- Evan Te - 2025年4月10日 星期四 上午10:20:00 [GMT+8]: +1 I'm an ASF member. I've watched 项目 grow from the ground up. It's a very active community with enormous impact. Given its maturity and scale as of now, I think it's also impressively sustainable. Would love to see ALC take the community to the next level.
- Ryan Lo - 2025年4月4日 星期五 晚上8:56:34 [GMT+8]: +1 I'm a committer at Apache YuniKorn & Submarine. Can't wait to see this come to life! Thanks WenChih (Ryan) Lo
- Luke Chen - 2025年4月4日 星期五 上午10:55:33 [GMT+8]: +1 I'm a PMC member of Apache Kafka. I'm completely supportive for this proposal! Thanks, Luke
- TingYao Huang - 2025年4月4日 星期五 上午11:41:54 [GMT+8]: +1 I'm a PMC member of Apache YuniKorn. It sounds good, I support this proposal. Thanks, Tingyao
- Kai-Hsun Chen - 2025年4月4日 星期五 凌晨4:00:18 [GMT+8]: +1 I'm a PMC member of Apache Submarine. I'm looking forward to having a formal Apache community in Taiwan. Thanks, Kai-Hsun Chen
- Wen-Chien Juan - 2025年4月4日 星期五 晚上9:35:12 [GMT+8]: +1 I'm a committer at Apache Submarine. Looking forward to the establishment of community in Taiwan. Thanks, Wen-Chien Juan
- Willem Jiang - 2025年4月4日 星期五 晚上7:30:24 [GMT+8]: +1 to start the local community in Taipei. Willem Jiang
- Wei-Chiu Chuang - 2025年4月3日 星期四 晚上11:45:57 [GMT+8]: +1 I'm an Apache Member and a member of several PMCs. So if it helps I can also join the Chapter's member team.
- Brandon Lin - 2025年4月4日 星期五 上午11:39:46 [GMT+8]: +1 Im a committer of Apache Submarine. Im very excited and looking forward to having a Apache event in Taiwan!
- Kevin Su - 2025年4月4日 星期五 下午5:38:30 [GMT+8]: +1 I'm an Apache Submarine PMC. I hope we can have a formal Apache community in Taipei. Thanks, Kevin
- Kuan Po Tseng - 2025年4月4日 星期五 上午11:18:11 [GMT+8]: +1 I'm an Apache YuniKorn Committer. It sounds like a great idea, and I'd love to see it happen. Best, Kuan-Po Tseng
- tison - 2025年4月3日 星期四 晚上11:57:43 [GMT+8]: +1 looks like a good proposal. Best, tison.
- Chung-En Lee - 2025年4月4日 星期五 中午12:40:24 [GMT+8]: +1 I'm a committer of Apache Ozone. I'd love to see this happen. Chung-En Lee
- Liu Xun - 2025年4月4日 星期五 上午11:39:42 [GMT+8]: +1 I'm an Apache Member and a member of several PMC & Committers. So if it helps I can also join the Chapter's member team.

你負責展現影響力 我們負責搭建舞台

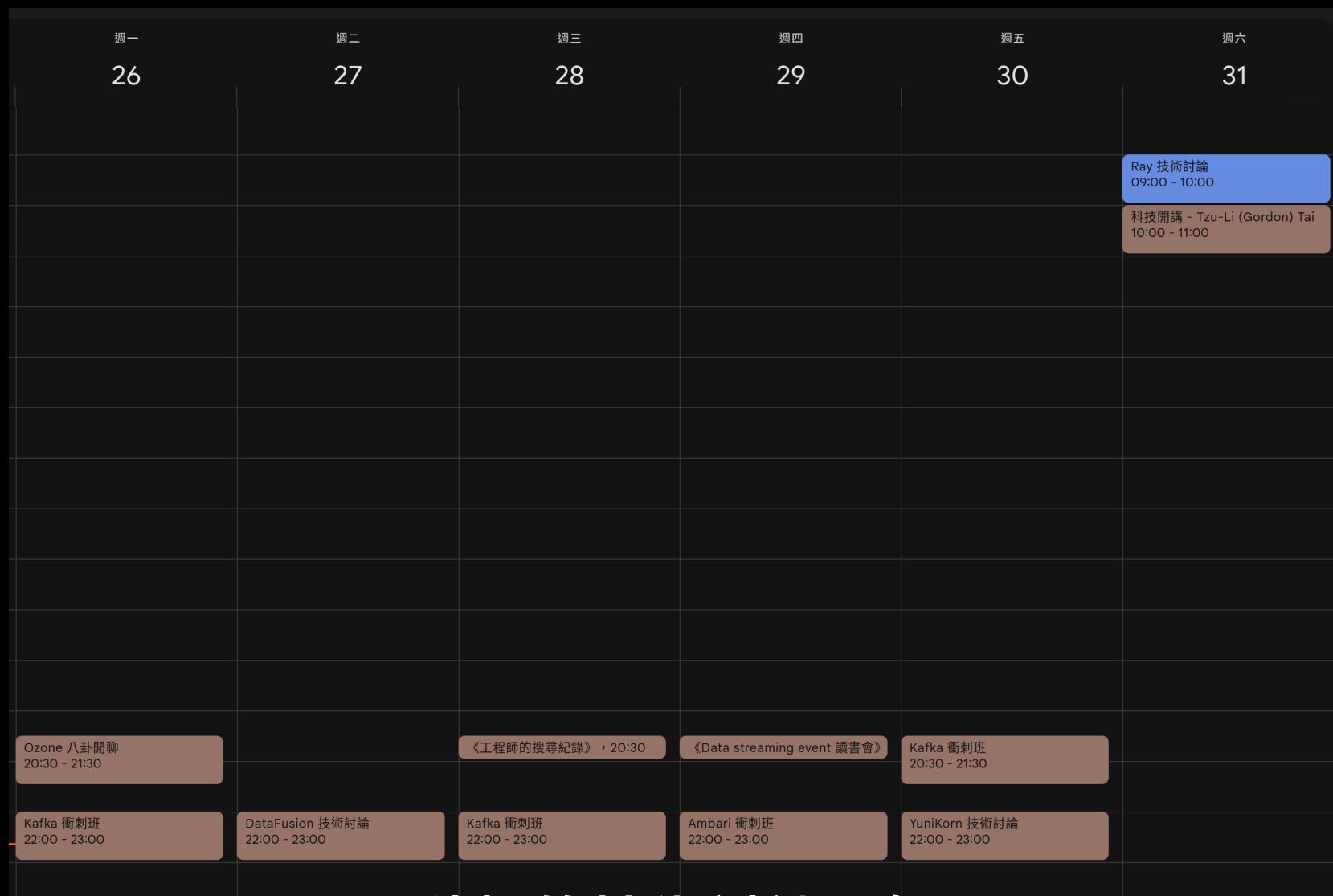
贊助設備



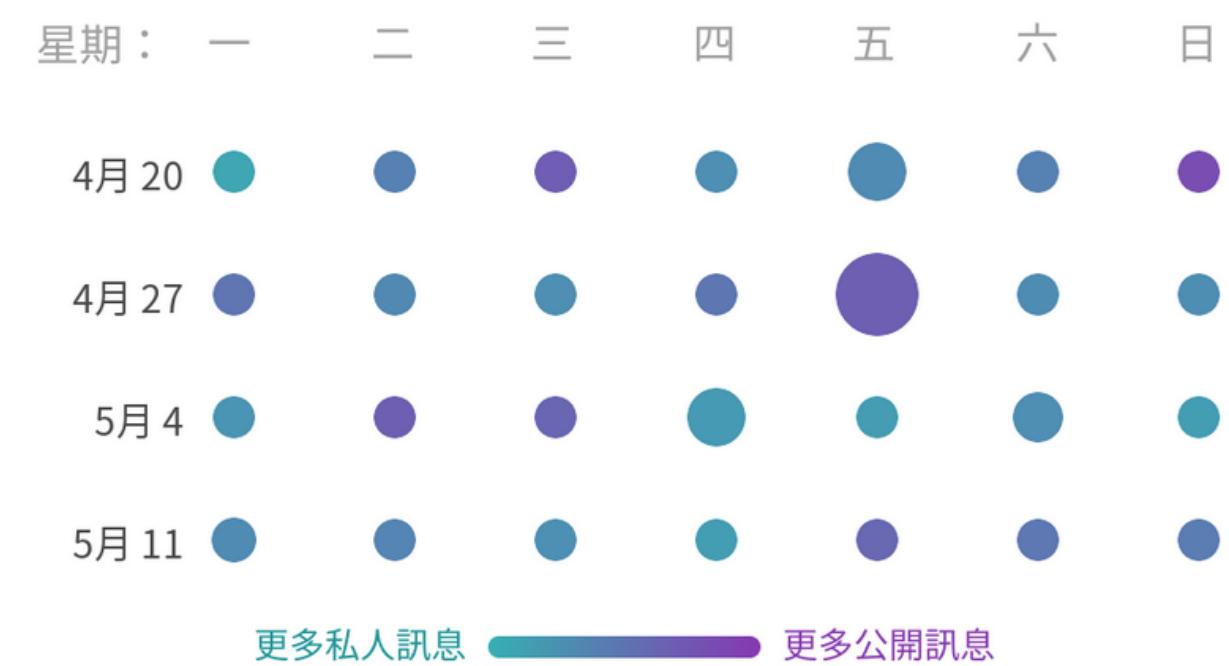
贊助交通費

日期	姓名	費用類別	金額	餐點
2025-05-02	**瑋	交通費	1350	菠菜饅頭營
2025-05-02	**佑	交通費	2655	菠菜饅頭營
2025-05-02	**呈	交通費	1350	菠菜饅頭營
2025-05-02	**成	交通費	1350	菠菜饅頭營

免費公開的線上技術討論



你的成員上週總共傳送了 3,210 則訊息 (比前一週增加 63 則)。其中，42% 位於公開頻道，2% 位於私人頻道，以及 56% 是私訊。你的成員也上傳了 127 個檔案 (比前一週增加 38 個)。



提供實驗環境



Jenkins

資訊主頁 >

+ 新增作業

‑ 建置歷程

⚙ 管理 Jenkins

☐ 我的視景

「源來適你」提供給社群夥伴貢獻開源的環境

贊助者：
郭靖 - 128 cores + 256 GB ram
蔡嘉平 - 128 cores + 256 GB ram

全部 +

S W 名稱 ▾

! ☁ kafka-e2e

圖示：小 中 大

重點回顧

① 參與開源除了當使用者之外，當貢獻者會更有價值

Beyond being a user, contributing to open source offers greater value.

② 當貢獻者對個人有許多好處，世界通用的技能、擴展人脈、國際認可的頭銜

Becoming a contributor offers many personal benefits, including globally applicable skills, an expanded network, and internationally recognized titles.

③ 公司參與開源貢獻有許多好處，處理安全漏洞、更容易招募人才、讓技術發展符合公司期待

Companies gain numerous advantages by contributing to open source, including the ability to address security vulnerabilities, more easily recruit talent, and ensure technical development aligns with company expectations.

④

開放任何拷問！

Chia-Ping Tsai