目录

[1：链上数据概述 2](#_Toc21119)

[2：The Graph是什么？ 2](#_Toc26992)

[3：The Graph是怎么工作？ 2](#_Toc4499)

[4：Graph node是怎么部署的？（通过部署Graph node来了解其中的运行机制） 2](#_Toc30201)

[5：智能合约的部署和子图的演示（以此说明子图的作用，具体是怎么从用GraphQL语言索引链上数据） 2](#_Toc1863)

# 1：链上数据概述

Web2行业的BI数据分析目前已经十分成熟，利用各种工具来分析已有数据的例子已经十分常见，例如零售商可以通过大数据分析预测下个月销售量大致为多少，避免库存积压。在Web3行业中链上数据分析目前也在不断发展，未来链上数据是一个十分巨大蛋糕。

但是其实目前要索引链上数据，其实效率上并不是特别高。1：要么自己部署节点（需要同步所有区块，耗时耗资源，对硬盘空间的要求也较高），2：要么需要使用到第三方节点托管服务，类似于Infura（这种方式效率十分低，索引数据速度特别慢）。

所以目前行业内急需比较高效的链上数据索引协议，The Graph是一种去中心化协议，从以太坊开始，用于索引和查询来自区块链的数据。它使得查询难以直接查询的数据成为可能。接下来我们具体来聊聊The Graph是什么，在链上索引数据起到什么样的作用，以及如何部署合约和子图演示。

# 2：The Graph是什么？

相信大家都知道CryptoKitties这款dApp，首先假设我们需要1）一个特定的以太坊账户拥有多少个CryptoKitty？2）一个特定的CryptoKitty是什么时候出生的?

这是因为合约暴露的balanceOf和getKitty方法直接支持这些读取模式。但是，其他问题更难以回答： 比如**在2018年1月至2月之间出生的CryptoKitties的所有者是谁？**

**要回答这个问题，您必须处理所有的Birth事件，然后为每个已出生的CryptoKitty调用ownerOf方法。**

**即使对于这个相对简单的问题，也要花费数小时甚至数天的时间。想要索引一个前端运行在浏览器的去中心化应用程序（DAPP）的区块链数据是非常困难的。**

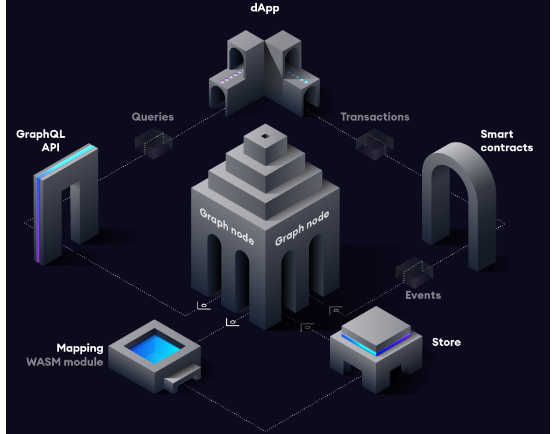
**基于以上问题，The Graph使用索引区块链数据的托管服务，这些索引（“子图”）可以用标准的GraphQL API索引数据。目前托管服务已经演变为具有相同功能的完全分散的协议。两者均受Graph Node的开源实现的支持。**

# 3：The Graph是怎么工作？

Graph根据子图描述（称为子图清单）来学习什么以及如何为以太坊数据建立索引。子图描述定义了子图感兴趣的智能合约，这些合约中要注意的事件以及如何将事件数据映射到Graph将存储在其数据库中的数据。

一旦编写了子图清单，就可以使用Graph CLI将定义存储在IPFS中，并告诉托管服务为该子图开始索引数据。

此图提供了有关子图清单部署后处理以太坊交易的数据流的更多详细信息：



流程遵循以下步骤：

1.去中心化应用程序(DAPP)通过智能合约上的交易将数据添加到以太坊中;

2.智能合约在处理交易时会发出一个或多个事件;

3.Graph节点持续扫描以太坊中的新区块以及可能用于您的子图的数据；

4.Graph节点在这些块中为您的子图查找以太坊事件，并运行您提供的映射处理程序。映射是一个WASM模块，用于创建或更新Graph Node响应以太坊事件而存储的数据实体；

5.去中心化应用程序(DAPP)使用节点的GraphQL端点向Graph节点查询从区块链索引的数据。Graph节点进而将GraphQL查询转换为对其底层数据存储区的查询，以便利用存储区的索引功能来获取此数据；

6.去中心化应用程序(DAPP)在前端UI中显示此数据，供最终用户使用在以太坊上发行新交易；

7.循环重复；

我们可以看到Graph node的重要性，接下来我们演示下如何部署智能合约和子图，通过GraphQL语言快速索引链上数据。