



企业级 可扩展图分析平台

DO IT WITH A GRAPH,
TIGERGRAPH



图数据库的兴起与未来

数据库发展概览： 从关系型数据库到非关系型数据库和图数据库

数据库与我们的商业管理和日常生活息息相关。它能有效追踪每笔商业交易，管理客户资料及库存信息，衍生相关经营报告。无论是普通电子商务平台还是社交网站翘楚（如Facebook），亦或是搜索引擎巨头（如Google），都离不开数据库的技术支持。

过去四十多年，数据库的发展随着用户需求转变而不断发展变化，各种计算机软硬件不断与时俱进，数据库的概念和功能不断变革创新，在这样的技术背景和时代契机下，TigerGraph应运而生。

近四十年来，由于使用了简便好用的SQL（结构化查询语言）及易懂的表格数据结构，加之行业生态环境较成熟以及相关工具的辅助（如Tableau和Microsoft Power BI），甲骨文、IBM DB2、微软SQL Server以及Teradata等关系型数据库几乎垄断了整个数据库市场。

过去十五年，数据库需求空前增长，传统的关系型数据库管理系统(RDBMSs)在处理大规模及复杂数据方面逐渐力不从心，无法满足各组织机构个性化的数据模式需求。

非关系型数据库借助其可扩展性及灵活应对大型数据库的能力逐渐普及开来，键值/宽列数据库（如Redis and Cassandra）和文件数据库（如MongoDB）都属于这类数据库。

然而，非关系型数据库同样有缺陷，因为它们的可扩展性是以牺牲其表现力为代价的。非关系型数据库使用的查询语言和应用程序接口与传统的关系型数据库使用的结构化查询语言相比，表现力要弱得多。它们不得不经常借助Java或Python编写的代码来完成那些用更高层级的语言（如SQL）写就的内容。

解决这些问题的关键就在于图的使用。图数据库不是什么新事物，它从现代计算诞生之初就存在了，但直到近年来其发展才渐趋成熟，并开始改变其既有地位。

从根本上说，图模式比关系模式要更具表现力。用关系型数据库设计的列表数据模式都能在图数据库里找到简单的匹配，但图数据库里有的却不一定能在关系型数据库里找到。每种结构化查询语言都能用图查询语言表达（例如TigerGraph的GSQL），图查询语言的表现力甚至更精准。很多成本较高的商业难题都可以使用图查询语言简单高效地解决，而这些都是关系型数据库无法企及的问题，比如客户聚类、PageRanking、欺诈团伙检测、供应链优化以及客户旅程分析等。

图数据库的优势主要在于其内在的数据架构。传统关系型数据库主要使用表格（由行和列组成）来存储和查询数据，需要磁盘或内存存储。而图数据库使用实体或关系（由顶点和边组成）来存储和查询数据。该架构差异赋予图数据库更优的性能。边能有效预关联实体，从而缩短查询时间。

预关联的重要性毋庸置疑。使用关系型数据库分析数据时，必须先建立一个由顶点和边组成的临时表格。使用图数据库分析，这些关联一目了然，可以跳过这一步直接开始分析。

列表数据架构对数据的查询和展示是二维的，而图数据库可以多维不受限地展示数据。例如，地理空间数据、组织层级、时间序列数据以及社会关系数据等，都能在图里自然成型。实际上，人脑思考和决策的方式更类似于图而不是表格。

图数据库的兴起与未来

为什么选图？为什么选TigerGraph？

尽管图数据库已经应用很多年了，但2017年TigerGraph的出现才算真正的变革。

构建商业级分布式图数据库极具挑战性，设计和实施分布式数据库系统面临诸多困难。与关系型数据库或者非关系型数据库相比，图数据库更需要随机存储以连接各个顶点来开展关联分析。

初代图数据库（如Neo4j）出现于十五年前。由于当时专为单机架构设计，其技术局限已然暴露，无法进行高性能的计算或支撑超大数据集。

以Hadoop和其它分布式非关系型数据库为代表的大数据时代来临前，图数据库主要应用在一些机构的小型项目上。TigerGraph在大数据时代背景下脱颖而出，突破性地开发了首个原生并行图数据库。目前全世界多家大型机构都在使用TigerGraph解决复杂的数据问题，这些机构包括世界十大金融机构中的七家、联合健康集团、捷豹路虎、Xandr、AT&T的一家分公司以及其他机构等等。

“原生”是指图数据库从零开始构建，为方便图数据处理而在数据存储、内存使用以及信息传递等方面进行专门的密集和优化设计。

“并行”是指图数据库不仅可分布于机器之上，还能同时利用每台机器的多内核。TigerGraph原生并行图构架将每一个顶点及边都作为数据平行存储和计算的单位，从而进行高性能的计算。

首个原生并行图数据库的问世对图数据行业产生了广泛而深远的影响。“原生”和“并行”这两项革新意味着一次范式转移，同时也使图数据库市场的发展更趋成熟。

常见问题

图用例

图数据库是一项普适的技术，广泛适用于具备高级数据分析和机器学习需求的各行业及用例，如：欺诈检测、反洗钱(AML)、实体解析、客户360、搜索引擎推荐、知识图谱、网络安全、供应链、物联网（IoT）、区块链及网络分析等。

TigerGraph的客户群有哪些？

我们很荣幸能与世界上很多业界巨头和有远见的大客户开展商业合作，这其中包括财富500强的部分大公司、创新型新公司以及一些中型企业。全球前十的金融机构有七间使用了我们的产品，全世界最大的电力公司以及多家制造业、通讯业、媒体、娱乐、零售业及电子商务公司也明智地选择了我们。TigerGraph还成为了医疗保险业最大的图数据供应商，被广泛应用在兼具临床研究和商业用途的医药公司上。

TigerGraph目前有哪些产品？

2017年以来，TigerGraph面向市场设计了一系列产品组合，其数据库管理系统既可安装在客户端服务器上，也能在AWS, Azure, and GCP三大云平台的云端运行。

TigerGraph 3.1 版本是一款原生图库，支持属性图模型和图查询语言。TigerGraph不仅为软件开发商提供了开放源代码工具箱，内置算法库和连接器，同时还配备了预置图数据模式和图查询功能的云端入门套件。

其本地（部署）企业版可提供高达50 GB的免费原始数据和免费云端相关服务。

TigerGraph在AWS, Azure及GCP云平台提供多云图数据库即服务，客户可根据自己的实际情况购买和使用这些云平台的TigerGraph云图数据库服务。

TigerGraph能轻松实现从关系型数据库到图数据库的无代码（no-code）转换，并通过可视化查询生成器，实现复杂图无代码分析，还支持Python语言。用户可通过pyTigerGraph 应用程序接口（API）连接机器学习架构，也可使用全局结构化查询语言（GSQL）进入机器学习内存数据库。通过高分布式平行计算，TigerGraph的数据存储架构可进行多跳（多层）查询。

图数据库市场到底有多大？

图数据库市场正在快速增长，而TigerGraph占据了市场较大的份额。

基于其卓越的性能和超群的表现力，我们预计可开发的图数据市值将高达840亿美元。它们主要来自以下五个板块：

1. 图数据库市场

图数据库本身就比面向文档的数据库（如MongoDB）要庞大得多。图数据库能解决很多文档数据库无法胜任的商业难题，并赋能更多高价值的商业应用。

2. 图赋能的机器学习和人工智能市场

机器学习（ML）和人工智能（AI）从根本上改变了很多行业。图分析可以多渠道赋能机器学习和人工智能发展。很多知名客户，如美国银行、摩根大通银行以及财捷集团都在使用TigerGraph加强其机器学习能力，并通过欺诈检测避免损失获取丰厚的投资回报。我们正在构建更多机器学习与人工智能内存数据库，并将与大型云端服务商的机器学习/人工智能构架相融合。这部分的市场潜力比分布式图数据库市场更大。

3. 图分析商业智能应用市场

商业智能应用可助力非专业人士用户管理、分析其关注的业务数据并使之可视化。关系型数据库商业智能市场前景广阔，是因为它对使用者的技术背景要求不高，数据使用不设限。从一开始TigerGraph就在GraphStudio visual SDK里设置了图数据可视化配置。2020年，TigerGraph推出了可视化查询生成器和无代码关系型数据库工具包。这些都是我们迈向商业智能的第一步。

我们认为该市场将远超分布式图数据库市场。TigerGraph是目前市面上唯一一个终端用户能开展实时交互式可视化图分析的分布式图数据库。这一优势使我们能在图商业智能应用方面和图可视化领域提供更优质的创新产品与服务。

4. 图赋能的垂直商业应用与数据/应用程序接口服务

目前我们已经开发了20套涵盖图数据模式（schema）、查询功能和样本数据的启动工具包，可迅速满足极具个性需求（例如供应链和客户实体管理）的相关应用，助力客户解决实际问题。但这些工具包远远不能满足市场需求。我们拥有越来越多欺诈检测、客户实体图分析、客户360及供应链优化用例，我们还将携手合作伙伴共同开发能提供一揽子/整套解决方案的专门应用，快速实现客户的投资回报。由于图数据可互相关联，图查询语言（GQL）功能强大，使用图数据库开发商业应用要容易得多。我们相信这部分的市场潜力也将远超分布式图数据库市场。

5. 基于图查询语言（GQL）的相关新应用的开发

TigerGraph是图查询语言（GQL）国际标准化委员会的成员之一，致力于在国际范围内建立一套强大、易用的属性图标准查询语言。图查询语言将成为数据科学家构建商业应用和开发机器学习与人工智能的新Python语言。就好像Java取代Cobol成为更受欢迎的应用工具一样，我们相信，从长远来看，图查询语言和图数据库将取代结构化查询语言及关系型数据库成为高级数据分析和新商业应用开发的首选。

这一板块的市场体量大得惊人，甚至比前四类图市场的体量总和还要大得多。看看全球总共有多少关系型数据库的供应商（例如甲骨文，微软和Snowflake）就不难想象这一点。TigerGraph也已朝着这个方向在努力。TigerGraph首席技术顾问Dr. Alin Deutsch以及他的技术团队在2021数据管理专业委员会（SIGMOD）论坛上首次发布了他们的原创成果，该论坛在数据库业界享有很高的声誉。TigerGraph用户正在使用其产品解决以前根本不可能解决的问题，或是那些因使用Teradata或Snowflake产品而进展缓慢的问题。随着图形查询语言的进一步开发，TigerGraph将会更频繁地出现在高级数据分析用例上，完成那些以往需借助高成本的关系型数据库才能完成的任务或是那些迟至今日仍然未竟的任务。



以上各板块也凸显了当下及未来几年TigerGraph产品开发的途径及公司的投资方向。作为业界唯一的可扩展企业级图数据库，并借助独一无二的可视化查询生成器在图技术商业应用方面独占鳌头。目前公司正全力开发图赋能的机器学习/人工智能应用和界面设计应用，并将在图平台，尤其是云空间加大创新和投入力度。

TigerGraph的未来发展蓝图？ 和最近的发展目标是什么？

我们的目标是在不到三年的时间成为年度经常性收入高达一亿美元的大公司，并随时准备上市。我们已经邀请国际四大会计师事务所入驻参与相关审计工作，并在公司内部启动了相关上市治理流程。

小结

图数据库是数据库技术发展史上的一次范式转移，2017年TigerGraph的出现重新定义了图数据库的作用与前景，图技术的发展也因此而渐趋成熟。

TigerGraph参与制定的图查询国际标准语言（ISO GQL）将为今后的商业应用开发提供一整套标准化的、高表现力的国际通用查询语言，大量数据库用户将从中受益。我们相信图查询语言必将成为数据分析的新Python语言，并完成许多目前关系型数据库和内部解决方案无法胜任的高级数据分析。

正如传统的商业应用工具（如Tableau，微软PowerBI和Looker）能在结构化查询语言非专业背景的用户中得以普及一样，图赋能的商业应用工具（如TigerGraph可视化查询生成器）也将在图数据库初始用户中得以广泛普及，使其能充分领略数据关联和图分析的无穷魅力。

附录

对SQL开发人员来说， 使用图数据库到底有多简单？

近年来，业界对图数据的关注度越来越高，大大加速了图查询市场的发展。创建GQL标准的初衷主要是为了满足数据科学家的需求，他们是否曾接触过SQL无关紧要。而TigerGraph意识到现有SQL开发人员未来将会是数目庞大的潜在用户，这也是为什么我们想要渐进式推动以图查询替代SQL查询的原因。

为此，TigerGraph设计了GSQL图查询语言，以最自然、变动最少的方式对SQL进行拓展，因为关系查询和图查询都能匹配数据查询。GSQL语言是制定GQL标准的参考语言之一，该标准最近甚至开始接纳GSQL语言影响下兼具SQL风格的句法。

图如何赋能机器学习 和人工智能发展？

今天的分析型图数据库（例如 TigerGraph）通过关联用户数据、更好呈现知识信息及实时获得更深层次问题的答案，提升用户业务水平。

这些优势同样惠及机器学习和人工智能领域。图数据库和图分析可以通过三种方式赋能人工智能：

- 使用图算法进行无监督学习
- 提取图模式、浓缩图特征
- 在图数据库里内置机器学习技术

附录

如果关系型数据库大供应商、云供应商或初创公司加入竞争， TigerGraph将如何保持业界领军地位？

目前，TigerGraph在市场竞争中已经处于优势地位。很多时候，我们的客户选择 TigerGraph是因为只有我们的产品才能够满足他们对高性能和可扩展性的需求。我们的产品性能比竞争对手强数百倍。

我们如何在竞争中完胜对手？公司创建的头五年，我们使用C++优化图性能、数据压缩和内存使用，从零开始自创了原生分布式图数据库。

我们是市场上唯一可扩展的图数据库。TigerGraph 产品推出以前，第一代 Neo4j 因为引领市场而对用户免费开放，因此图市场没有太多竞争。TigerGraph产品进入公众视野后，一些初创公司也竞相效仿。他们仍在积极摸索中，但这些初创公司在产品成熟度、性能及可扩展性方面落后我们数年时间。此外，有些公司还在使用我们曾考虑采用但最终摒弃的设计架构，例如稀疏矩阵计算。

今天，我们本地部署的主要竞争对手是 Neo4j，而云端的主要竞争对手是 Amazon Neptune，但它甚至比 Neo4j 运行还要缓慢。有关性能、可扩展性和功能方面的详细对比，请参阅[针对Neo4j和Amazon Neptune的详细图数据库购买指南](#)。

竞争对手分析：Neo4j

缺点：从更高的技术层面上看，Neo4j 最大的缺点是它的单机架构，其赋能用例非常有限。

优点：Neo4j最大的优势在于它作为图领域的先行者，已经做了20年的营销和社区建设，在业界享有盛誉。然而，TigerGraph 正策略性地培育开发者用户社区：免费的云产品、免费的本地企业版本以及由我们的开发者公关团队管理的众多免费在线论坛和活动。再加上我们无与伦比的性能和可扩展性，TigerGraph在最近的 Forrester Wave 报告显示中已经成为公认的行业领导者，在用户社区建设以及与企业规模部署相关的其他总共9 项标准中均得了 5 分（满分 5 分）。

Neo4j 4.0 Fabric 架构表现如何？我们分析了 Neo4j 4.0 Fabric 架构，将其与 TigerGraph 进行比较，探讨了图数据库企业的可扩展性国际标准。免费下载《图数据库对比指南：TigerGraph vs Neo4j》，指导图数据库选型。

竞争对手分析：Amazon Neptune

缺点：Amazon Neptune是几年前一家 10 人组成的初创公司 Blazegraph 被收购后再度推出的改版。其最大的缺点是单机架构性能，运行速度甚至比 Neo4j 还要缓慢。从根本上说，Amazon Neptune 实际上是一个 RDF 三重存储图，可以配置用作属性图，支持Gremlin/Tinkerpop 语言子集，这是一种越来越少人用的Apache 开源框架。Neptune不支持OLAP分析。

优点：Amazon Neptune最大的优势在于它因 AWS 品牌而广为人知。从好的方面来说，这有助于用户增加对图的了解，并能带动更多人尝试在云平台使用图，这对 TigerGraph大有裨益。当用户熟悉图之后，他们很容易发现只有 TigerGraph 产品才具备那些大型项目所需的高性能和可扩展性。我们的很多客户有类似经历，先试用了 Neptune，但最终选择了TigerGraph。

科技巨头

——微软、谷歌、Facebook、IBM 和Oracle

微软在云端有 Cosmos DB，它配置了较为初级的图应用程序接口（API），但它又不是图数据库。因此，我们不将其视为竞争对手，而是寻求与 Microsoft Azure 更紧密的合作，以充实他们的图产品。

附录

它不支持适用于多功能商业应用开发的高级查询语言。与 Microsoft Azure 一样，我们将寻求与 Google GCP 在图方面的合作。

Facebook支持的Apache Giraph在最近几年没有进一步的突破，它只支持离线图分析（无更新，只读，无查询语言，只有Java应用程序接口）。Facebook 拥有构建在多套MySQL系统之上的图内部解决方案，但只针对专门应用程序，使用范围狭窄，不支持其他应用程序。

Oracle在其RDBMS上支持图运行，但在性能方面，它无法与TigerGraph这样的原生图数据库竞争。除了性能之外，Oracle的图是只读的——更新必须通过RDBMS表单才能实现。

几年前，IBM发布了基于JanusGraph的IBM Graph Cloud。由于其性能缓慢，我们并未将IBM图视为竞争对手。与Microsoft Azure类似，我们将寻求与IBM Cloud开展合作。



TigerGraph的核心优势

我们的产品是付费的，我们的系统是从零开始构建的，用于从图存储到图处理引擎的并行操作。

TigerGraph 的核心优势包括：

- 得益于原生图和从底层设计的高性能
- 因其分布式架构而具备的可扩展性
- 易用、功能强大且图灵完备的查询语言。这意味着 TigerGraph业务团队和客户几天之内就能完成概念验证或试运行从而彰显其业务价值，而不是像使用其他系统那样需要耗时数周或数月。
- 图可视化和商业智能（BI）。以前，由于缺乏高性能的图数据库，市面上没能出现好用的图可视化和商业智能工具。TigerGraph 是唯一支持实时、交互式图商业智能和图可视化任务的引擎。因此，TigerGraph 通过无代码/低代码工具（例如 GraphStudio 和可视化查询生成器）为图领域带来了创新。

我们的竞争对手所面临的障碍

构建分布式图数据库耗时耗力，需要丰富的专业知识。以 Neo4j 为例。TigerGraph于2017 年未进入公众视野之后，Neo4j 曾试图解决其可扩展性问题，因为它必须抗衡TigerGraph，而TigerGraph是市面上唯一的分布式图数据库。Neo4j 的第一步举措是利用 Spark框架中的分布式功能来分析离线图，但该尝试以失败告终。

Neo4j 也曾尝试使用多台机器来解决可扩展性问题，但 Neo4j 并没有设计真正的分布式图系统，而是尝试使用一种名为 Neo4j Fabric 的解决路径，通过一台机器作为协调器，提供技术脚本与独立的 Neo4j 机器进行对话，开发人员必须在每台独立机器上都手动创建图模式和图查询（可查询买家指南了解更多信息）。尽管 Neo4j 在过去的 20 年里一直在研究图，但他们依旧很难创建分布式图数据库。

附录

除面临构建分布式图数据库的技术挑战外，我们的竞争对手还面临以下困难：

1. 即使他们成功赶上了我们，也不会比 TigerGraph性能更优越。

2. TigerGraph已通过用户的实战测试，并将在现有基础上继续添加企业功能（例如 MultiGraph）和各类不同级别的数据访问控制。选择数据库时，客户不仅在意性能，还会考量其功能、稳定性、安全性和许多其他需要耗时数年才能配备的新应用。

3. 作为行业先锋，TigerGraph 是图查询语言（GQL）国际标准委员会的成员，引领国际查询标准的创建。作为该委员会的新兴力量，我们能够提前了解GQL 标准可能包含的内容及其影响。

4. 因为只有TigerGraph具备企业所需的高性能和可扩展性，包括Intel、HP Enterprise、Dell和Xilinx在内的大型硬件供应商正在与TigerGraph合作，加快TigerGraph在其硬件上的运行速度。这意味着TigerGraph可以在图领域继续拓展技术创新。

5. TigerGraph 在无代码/低代码应用程序开发、图商业智能和图可视化方面遥遥领先。我们将继续加大创新和投入力度，提高图数据库的利用率。

6. 图+机器学习。我们正在系统中添加更多的图机器学习和人工智能功能，以便数据科学家能轻松快捷地使用图数据库。同样，得益于我们的高性能，只有TigerGraph 才能真正展现图+机器学习的优势。



关于TigerGraph

TigerGraph是一个基于数据关联的高级分析和机器学习平台。TigerGraph是业内首个也是唯一一个分布式原生图数据库，在欺诈检测、反洗钱（AML）、实体解析、客户360°、搜索引擎推荐、知识图谱、网络安全、供应链、物联网(IoT)和网络分析等应用场景效果显著。更多信息，请访问中国官网：www.tigergraph.com.cn



© TigerGraph, Inc. 2022 All Rights Reserved. © TigerGraph, Inc. 2022 版权所有。



TigerGraph服务号



TigerGraph订阅号