

让数据改变工作与生活

BONC 东方国信

CirroData Graph 产品介绍



分布式数据库研发中心

2021

图数据库是什么

CirroData Graph介绍

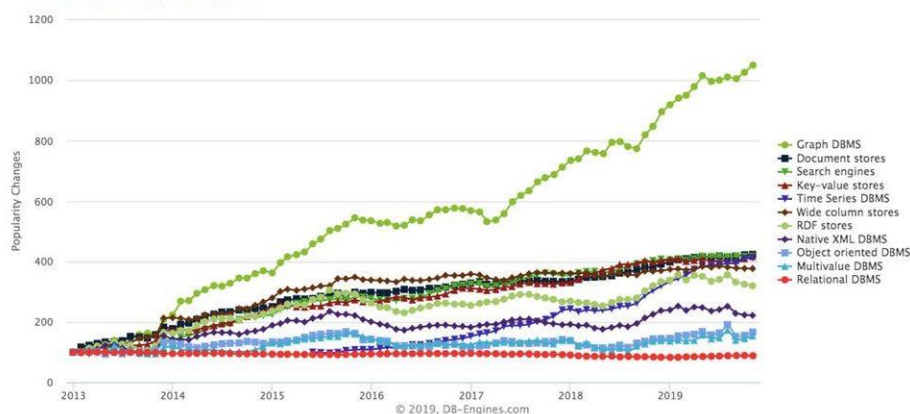
为什么是图数据库

CirroData Graph架构与性能

图数据库应用场景

图数据库选型

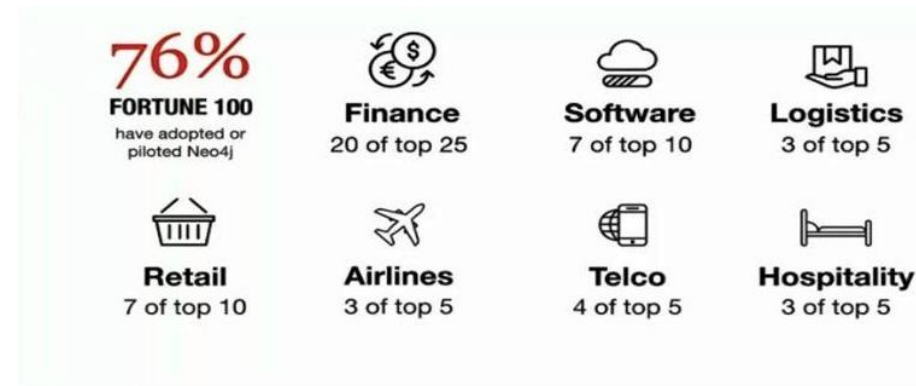
急速增长中的图数据库



数据库软件评比网站DB-Engines的一个走势图。图数据库的增速从2013年开始一直是最快增长的数据库分类，像数据库顶尖大会如VLDB也是图数据库的相关论文数能排到第二。

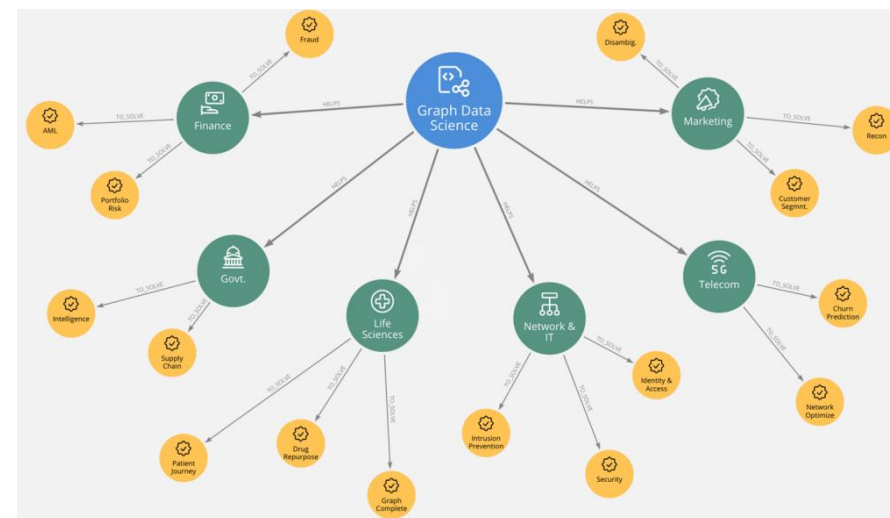
Operational 图数据库	RDF图数据库	多模式图数据	分析及大图数据库
Neo4j	AllegroGraph	Azure Cosmos DB	Apache Giraph
Titan	Blazegraph	ArangoDB	Turi
JanusGraph	Virtuoso	Sqrrl	TigerGraph
OrientDB	Stardog	RedisGraph	Cayley
			Dgraph

图数据库增长非常快，并且还有很多细分领域，标红的是研究比较深入或者用得比较多的。



<https://syncdreview.com/2018/10/04/graphconnect-2018-linking-ai-and-graph-db>

目前世界一百强企业使用图数据库的比例。金融，软件，零售,物流，电信，医疗行业。



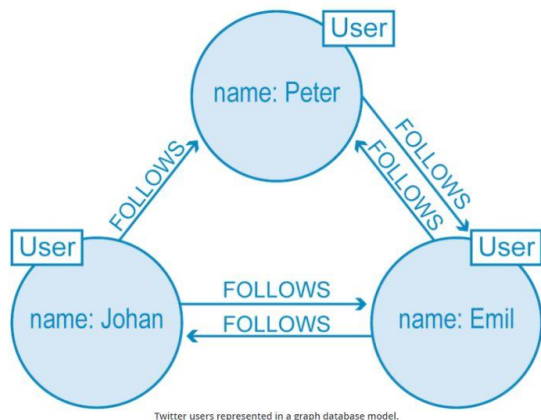
图数据库应用一图胜过千言万语

Q: 什么是图数据库?

A: 图数据库是图数据库管理系统的简称, 使用图形化的模型进行增删改查的数据库。图数据库一般用于OLTP系统中, 提供在线事务处理能力。与图数据库对应的是图计算引擎, 一般用于OLAP系统中, 提供基于图的大数据分析能力。

Q: 图数据库如何表达数据? 或者其建模方式

A: 图数据库使用图模型来操作数据。目前使用的图模型有3种, 分别是属性图 (Property Graph)、资源描述框架 (RDF) 三元组和超图 (HyperGraph)。现在较为知名的图数据库主要是基于属性图, 更确切得说是带标签的属性图 (Labeled-Property Graph), 当然标签不是必须的。下面是使用带标签的属性图的Twitter用户关系。



图数据库：存储实体与实体之间的关联关系

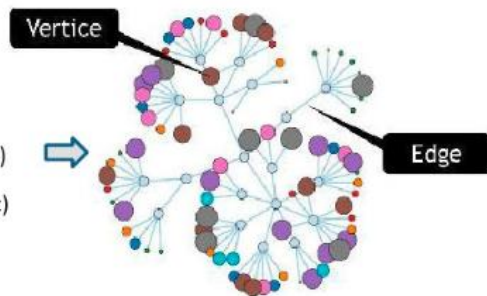


现实世界

$$G = (V, E)$$

Vertex (Node, Entity, Object)

Edge (Relationship, Link, Arc)



图数据库

为什么需要图数据库

Q: 为什么需要图数据库，相比关系型数据库等有什么优势？

A: 因为关系型数据库不擅长处理数据之间的关联关系。尤其是数据量比较大时。

右图是一个使用关系数据库模型实现的商品交易的例子。如果要做商品推荐相关的查询时，可以发现其中一些查询是低效的(例如，“客户购买了什么产品?”)，而有些查询可能是无法完成的(例如，“哪些客户购买了该产品?”)。

User					
UserID	User	Address	Phone	Email	Alternate
1	Alice	123 Foo St.	12345678	alice@example.org	alice@neo4j.org
2	Bob	456 Bar Ave.		bob@example.org	
...
99	Zach	99 South St.		zach@example.org	

Order	
OrderID	UserID
1234	1
5678	1
...	...
5588	99

LineItem		
OrderID	ProductID	Quantity
1234	765	2
1234	987	1
...
5588	765	1

Product		
ProductID	Description	Handling
321	strawberry ice cream	freezer
765	potatoes	
...	...	
987	dried spaghetti	



图数据库是一种新型的NoSQL数据库吗？

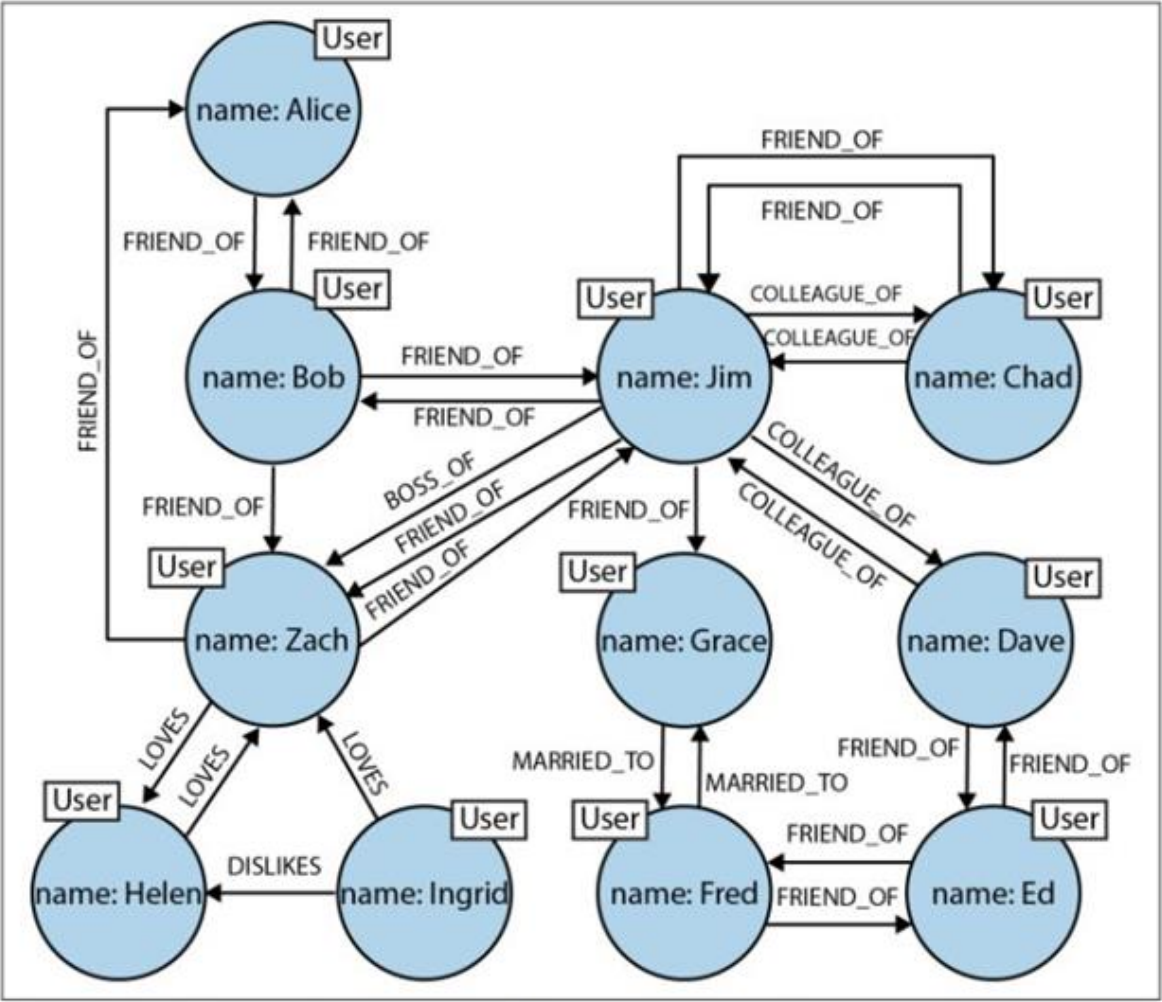
Q: 有关图的性能测试数据？

A: 举最经典的社交网络中查询的性能作为对比。

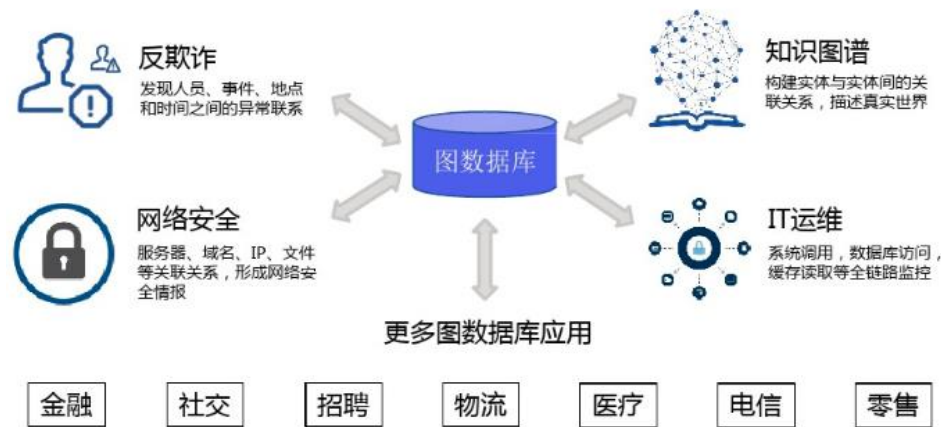
右图为一个社交网络，图中包括了朋友、同事、夫妻和恋人等多种关系。有人曾做过一个测试：在一个包含100w人，每人约有50个朋友的社交网络中找到最大深度为5的朋友的朋友。下图为图数据库Neo4J和关系型数据库在寻找扩展朋友时的性能对比。

Depth	RDBMS execution time(s)	Neo4j execution time(s)	Records returned
2	0.016	0.01	~2500
3	30.267	0.168	~110,000
4	1543.505	1.359	~600,000
5	Unfinished	2.132	~800,000

从图中可以发现，深度为2时（即朋友的朋友），两种数据库性能相差不是很明显；深度为3时，很明显，关系型数据库的响应时间30s，已经变得不可接受了；深度到4时，关系数据库需要近半个小时才能返回结果，已经枉称在线数据处理系统了；深度到5时，关系型数据库已经掉入深渊。而对于图数据库Neo4J，深度从3到5，其响应时间均在3秒以内。可以看出，对于图数据库来说，数据量越大，越复杂的关联查询，约有利于体现其优势。从深度为4/5的查询结果我们可以看出，图数据库返回了整个社交网络一半以上的人数。



Easily modeling friends, colleagues, workers, and (unrequited) lovers in a graph



一、社交网络应用

社交是人与人之间的连接，以图数据模型为内在的图数据库天生适用于明显的以联系为中心的领域。在社交网络中使用图数据库可以方便得识别人/群组和他们交流的事物之间的直接或间接的联系，使用户能够高效地对其他人或事物进行打分、评论、发现彼此存在的关系和共同关系的事情。可以更加直观得了解社交网络中人与人之间如何互动、如何关联、如何以群组的形式来做事情或选择。

社交网络是最基础的图模型，在此基础上可以叠加更多的内容，比如个人的喜好、购买过的物品、日常的生活方式等，从而演化出更高级的图数据库应用模式，比如实时推荐系统。

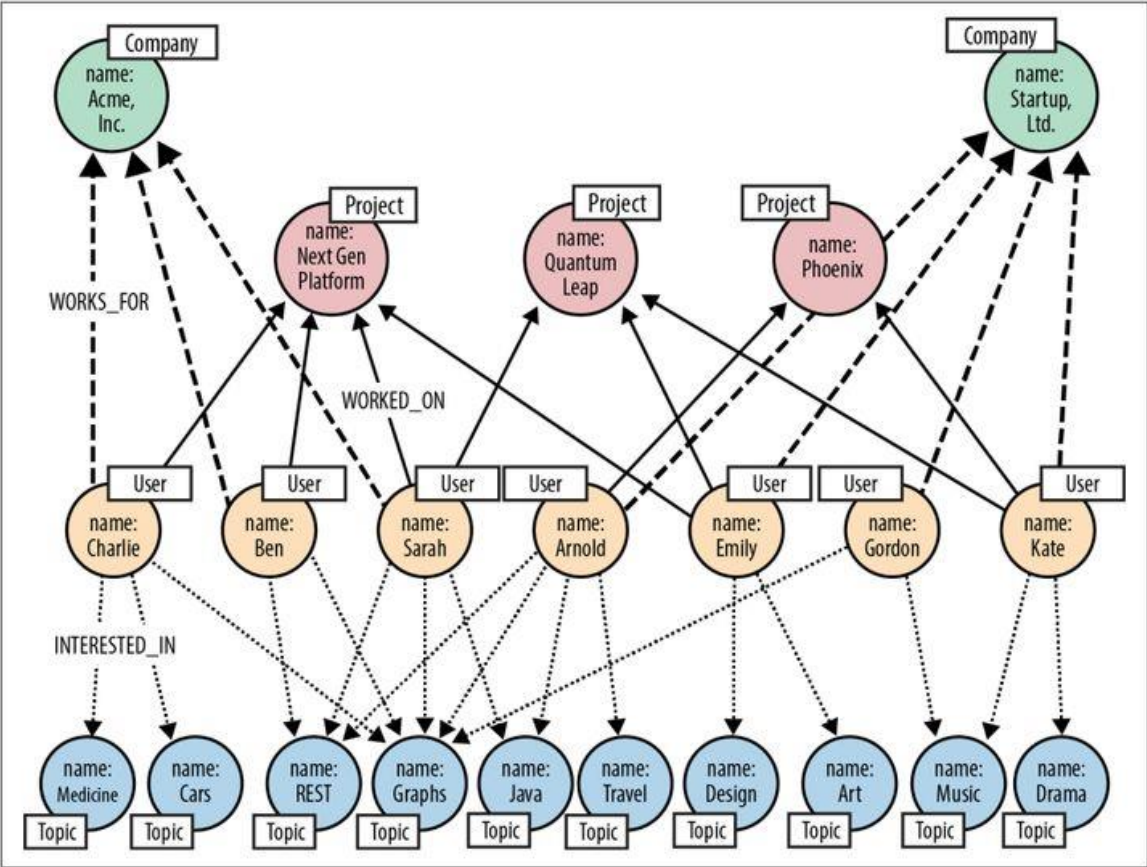


Figure 5-1. Sample of the Talent.net social network

二、地理空间管理

地理空间类的应用程序包括公路网、铁路网等，地理空间操作依赖于特定的数据结构，简单的加权带方向的联系，复杂到空间索引如R树。和索引一样，这些数据结构天生就以图的形式呈现，尤其是层级结构，非常适合图数据库。

总的来说，通信、物流、旅游已经路由计算相关领域的地理空间应用经常会使用图数据库。

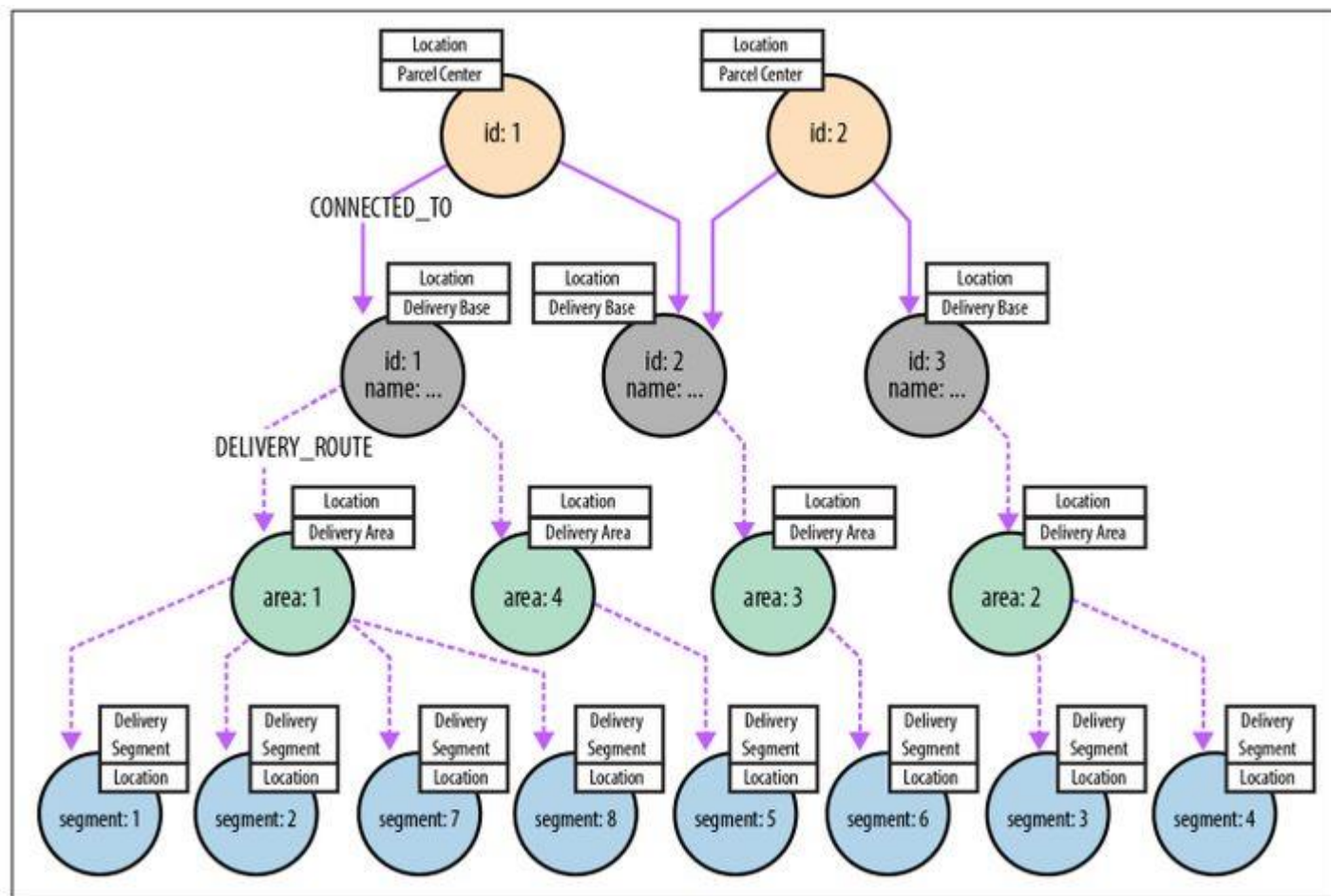


Figure 5-12. Structure of the delivery network for Period 1

三、实时推荐

在零售、招聘、情绪分析、搜索和知识管理领域，社交网络和推荐引擎可以提供关键的差异化能力，有很多种办法可以实现推荐，但使用图数据库在实时性和效率上有其特有的优势。推荐算法在人和事物之间建立联系，而联系建立的基础是用户的行为，比如购买、生产、消费、打分或评论有关资源等行为。推荐引擎可以识别出某些资源会吸引特定个人或群体，或者某些个人或群体可能对特定资源感兴趣。

一个有效的推荐依赖于对事物之间关联的理解，同时也依赖于这些关联的质量和强度，而属性图是所有这些关系密切、关联紧密的数据结构的最佳表达方式。用图数据库存储和查询这些数据使得应用程序可以为最终用户呈现实时结果，反映数据最新的变化，而不是返回给用户那些预计算的状态结果。

四、主数据管理 (Master Data Management)

在企业或组织中，主数据管理(MDM)包括的数据涉及用户、客户、产品、供应商、部门、区域、站点、成本中心和业务单元等。这些数据来源可能是多种多样的，MDM用来识别、清洗、存储和管理这些数据。其关键问题包括谁组织结构的变化、企业合并和业务规则的变化来管理这些变化；融合新的数据源，用外部源数据补充已有的数据；解决报告需求、鉴定需求和商业智能客户的需求；当数据的值和模式变化时对数据进行版本管理。图数据库的数据模型高效匹配MDM的快速演变和不断变化的业务需求。

五、网络和数据中心管理

图数据库已经成功地使用在了电信、网络管理和分析、云平台管理、数据中心和IT资产管理以及网络影响分析等领域。在这些领域里，他们将影响分析和问题解决的时间从数天数小时减少到了分钟级甚至秒级。面对不断变化的网络模式，图数据库的性能和灵活性都是它适合这些领域应用的重要因素。

六、授权和访问控制

图数据库可以存储那些复杂的、高度关联的、跨越数十亿参与者和资源的访问控制结构。尤其适用于内容管理、联合授权服务、社交网络偏好已经软件服务化提供。将这些系统从关系型数据库切换到图数据库后，性能从分钟级提升到毫秒级。

上面仅列举了部分例子，除此之外，图数据库产品还广泛用在金融和保险行业反欺诈、风控，电商和社交类产品防机器人作弊等领域。

总结

图数据库是一种越来越受关注的新型NoSQL数据库，图数据库研究和实践相关的论文占据了数据库领域顶尖会议和期刊中很重要的部分。目前国内有些企业也已推出了自己的图数据库系统，有自主研发的，也有基于开源图数据进行优化的。作为数据库从业人员，有必要对图数据库有最基本的了解。

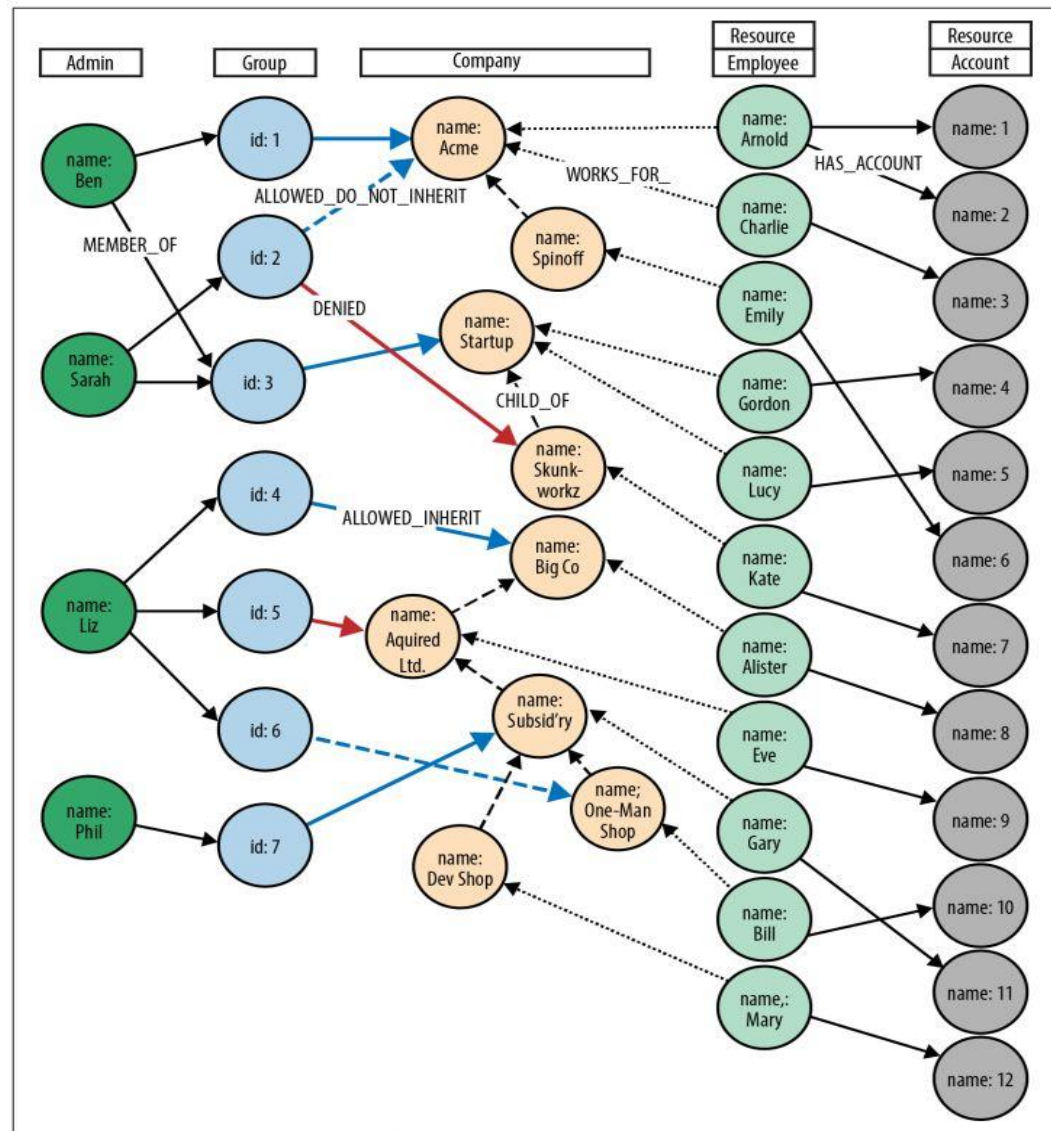


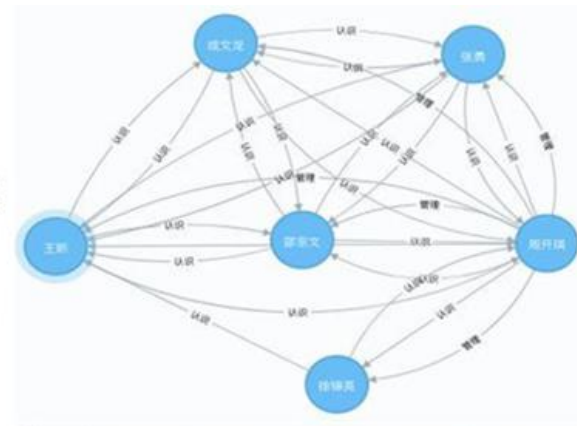
Figure 5-8. Access control graph

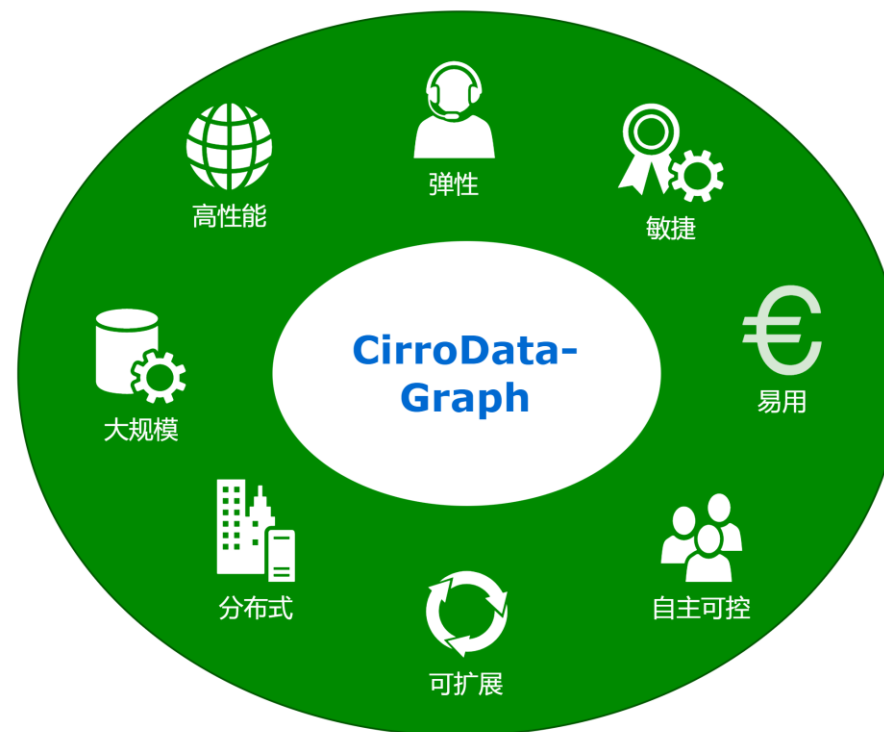
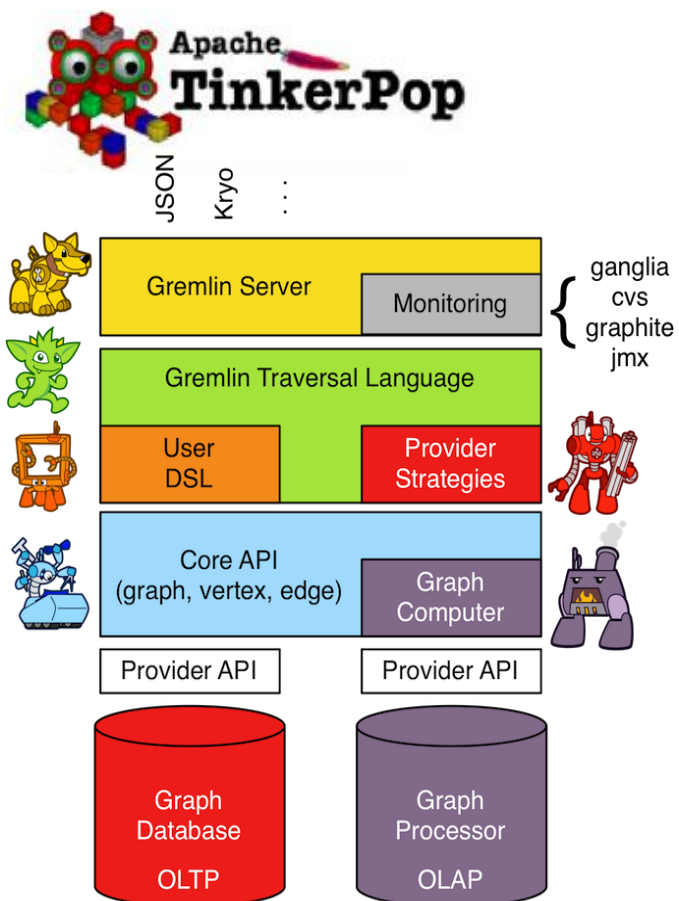
CirroData Graph的主要特点包括:

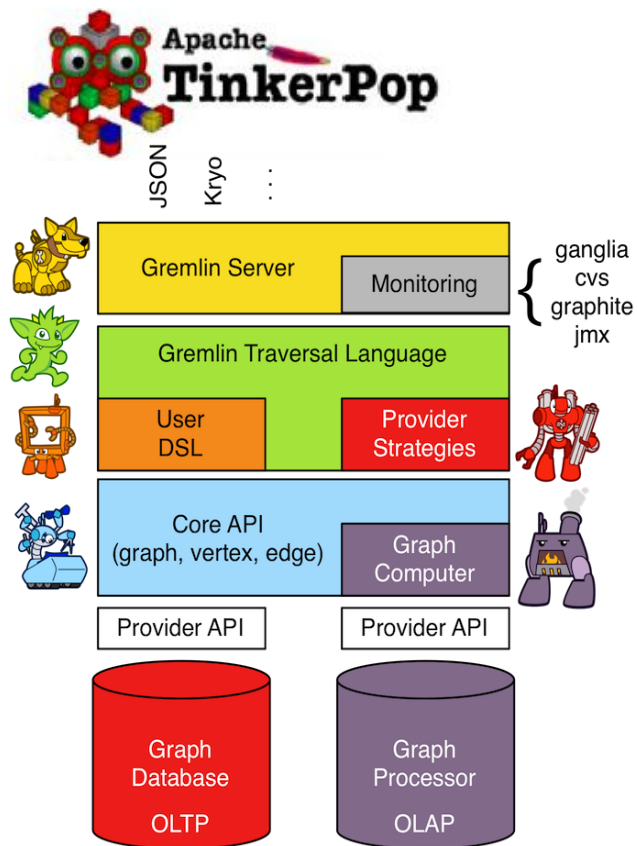
- ## 兼具直观性的图数据库

一冬胜过千言万语

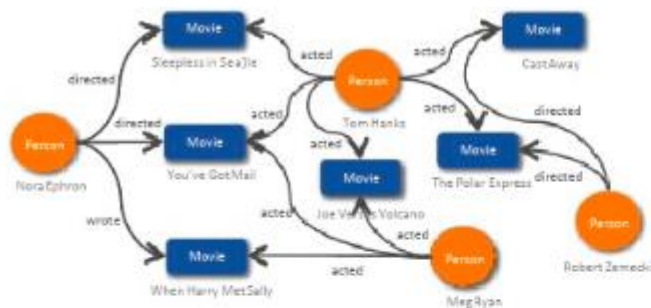
比起传统的信息存储和组织模式，图数据库能够很清晰揭示复杂的模式，尤其在错综复杂的社交、物流、金融风控行业效果更为明显。



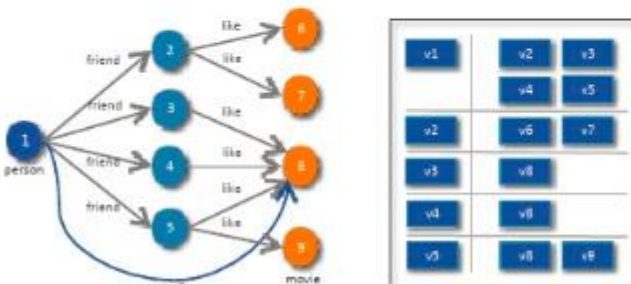




易建模



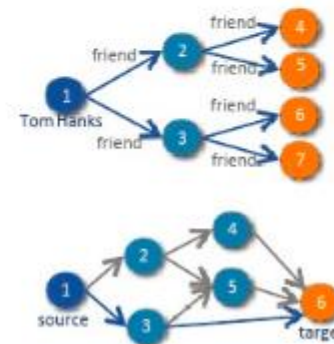
高效关联分析



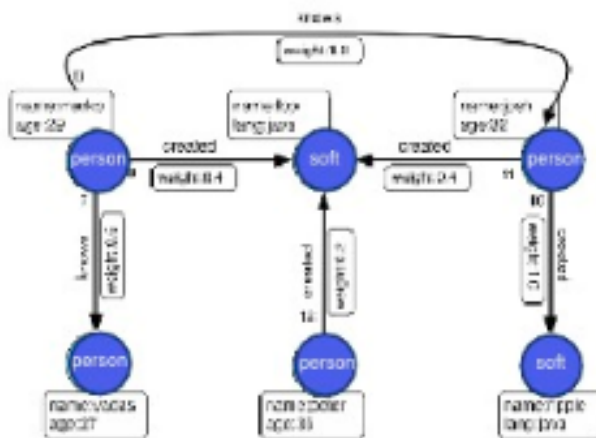
大规模



灵活查询语言



1. 概念模型



- **Property Graph**
- RDF

2. 存储模型

A: (B,D)
B: (C,D)
C: (E)
D: (B,C,E)
E: (A,C)
F: (D,E)

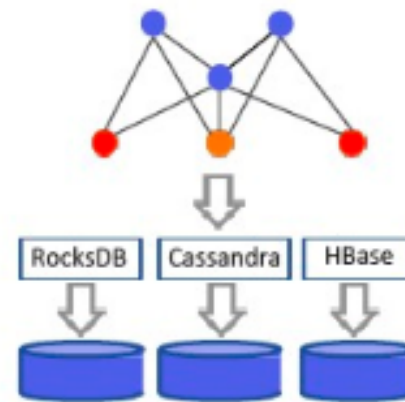
- **Adjacency List**
- Adjacency Matrix

3. 查询语言

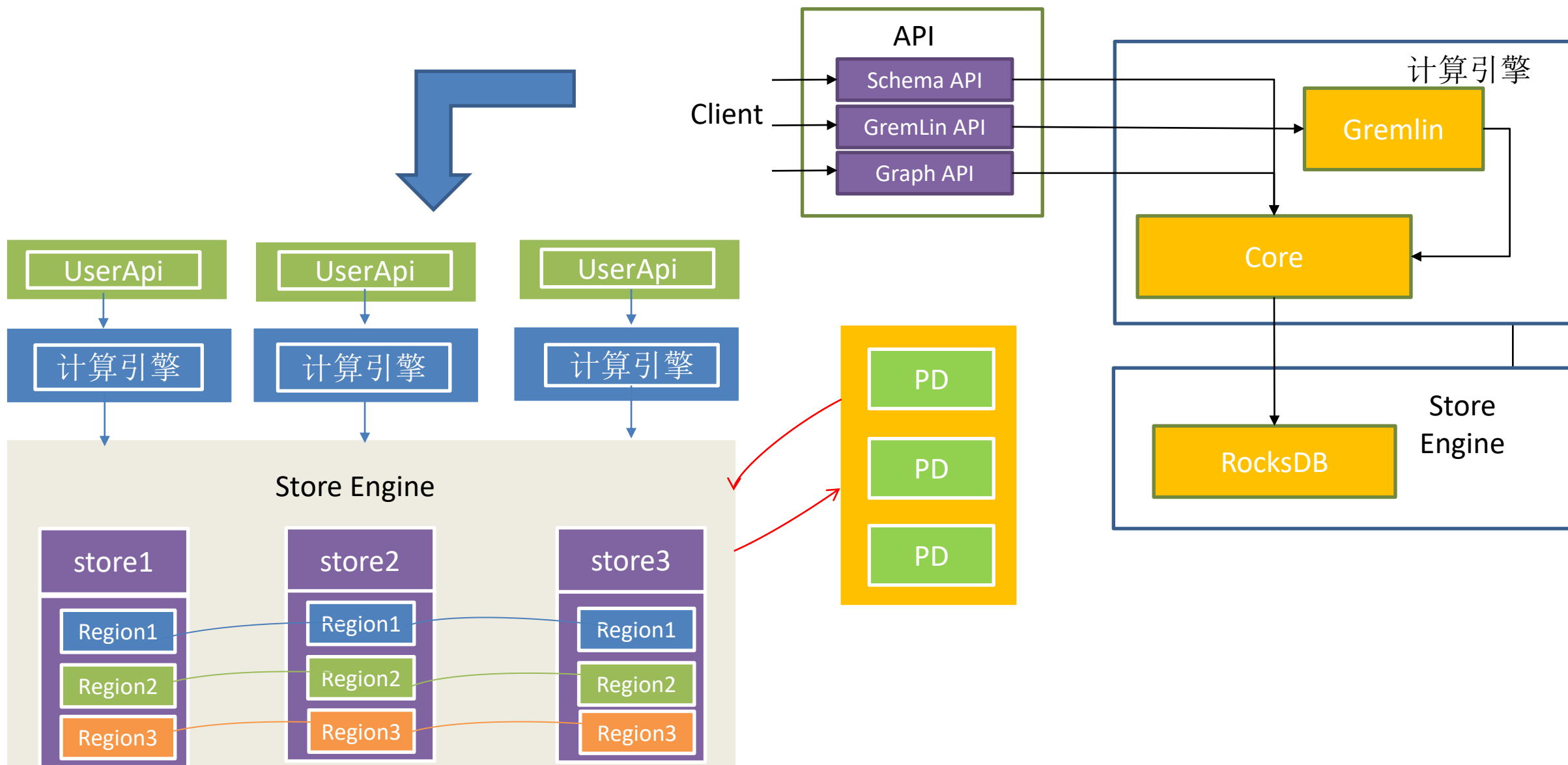
```
g.V().has('name','gremlin')  
  .out('knows')  
  .out('knows')  
  .values('name')
```

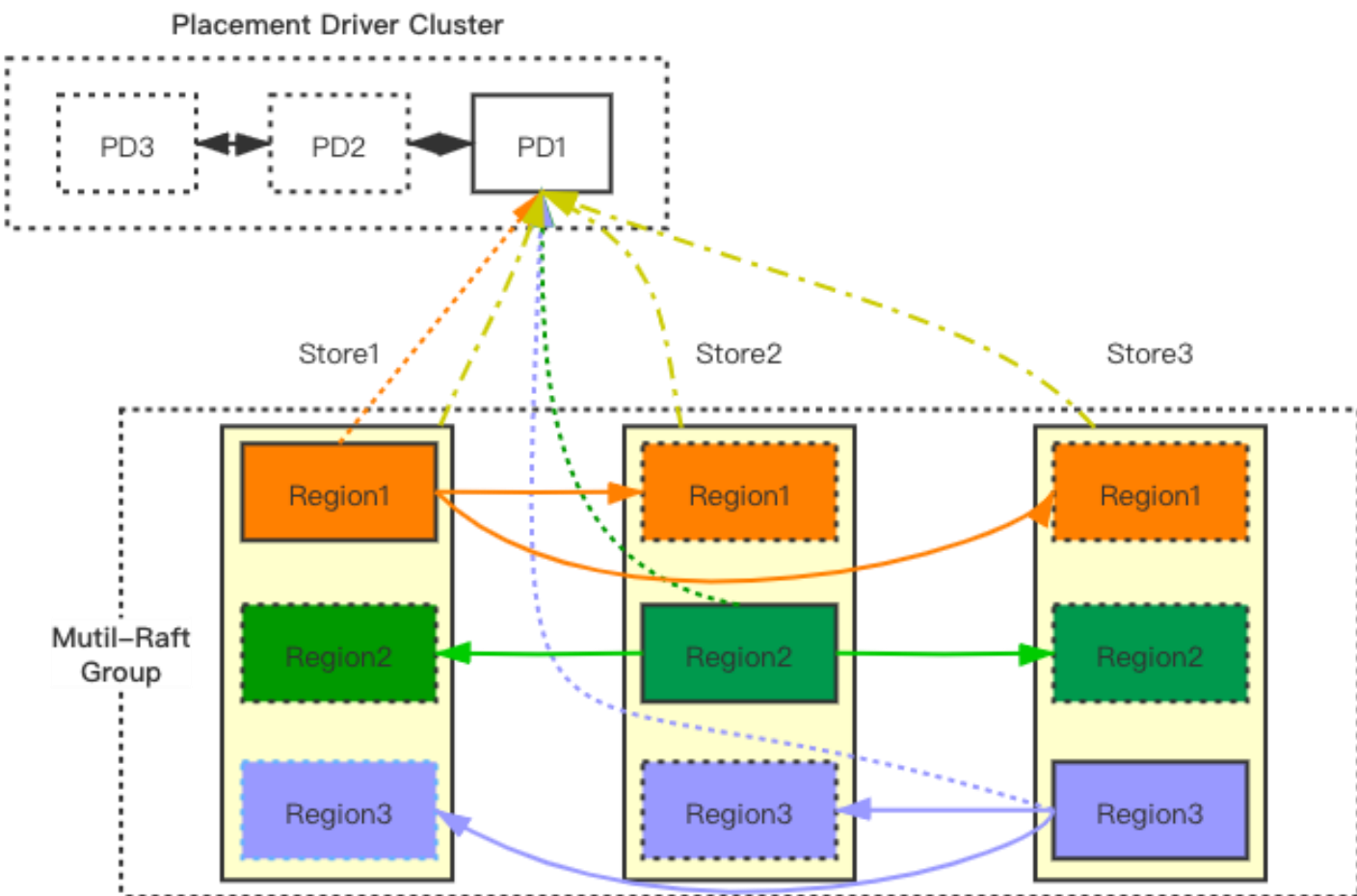
- **Gremlin**
- Cypher
- SPARQL

4. 持久化方案



- **Non-Na3ve Storage**
- Naive Storage





图中Store1、2、3中的region leader分别为region1、2、3。

Store向pd发送心跳

region leader向pd发送心跳。

region leader向follower发送日志。

Placement Driver

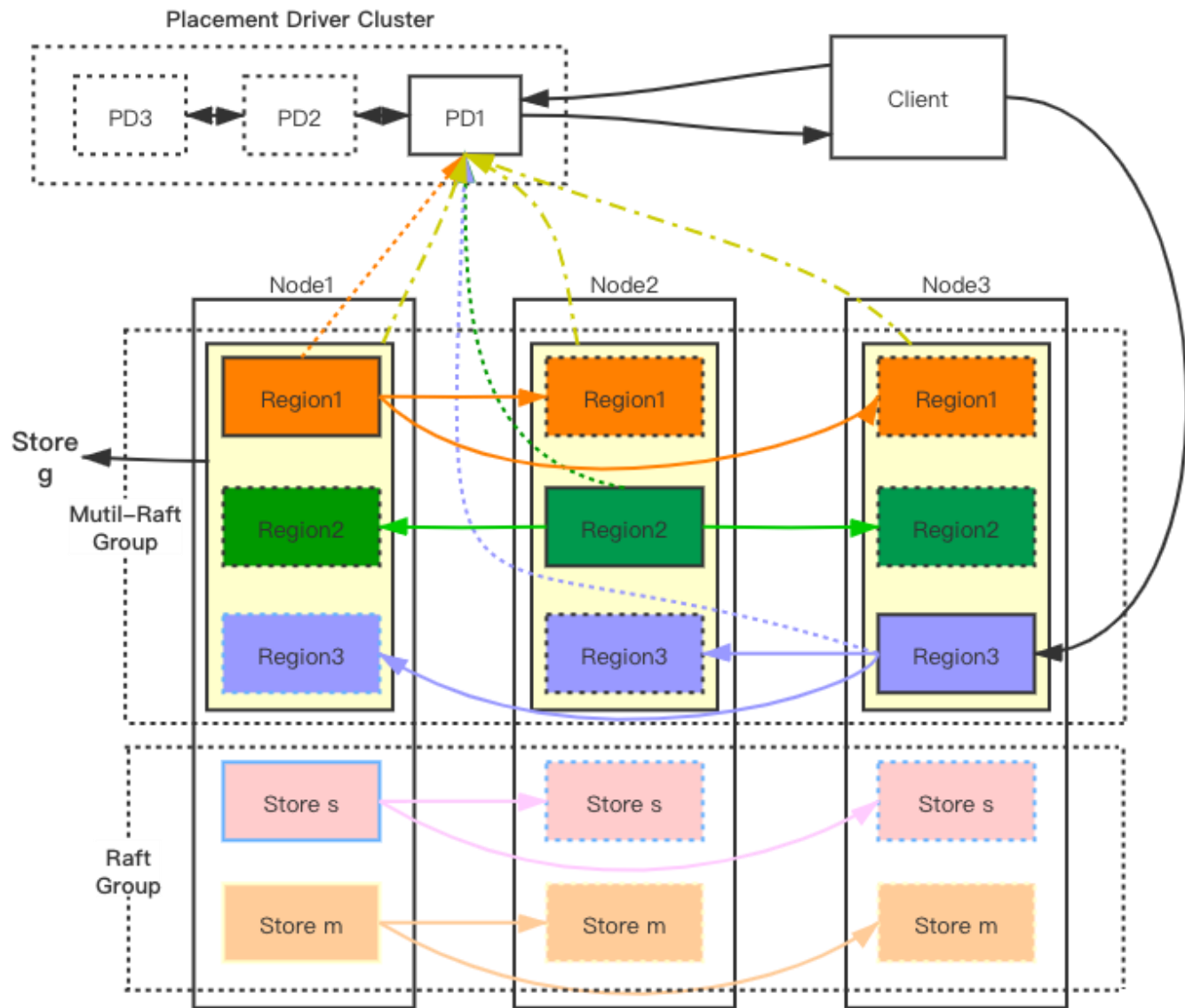
简称PD，PD服务需要单独部署，主要负责整个集群的管理调度，Region ID生成等。

Store

一个节点一个Store，可看作Region的容器，储存多个分片的数据。

Region

数据存储的最小单元，每个Region存在多副本，存在不同的Store中，一起组成Raft Group。

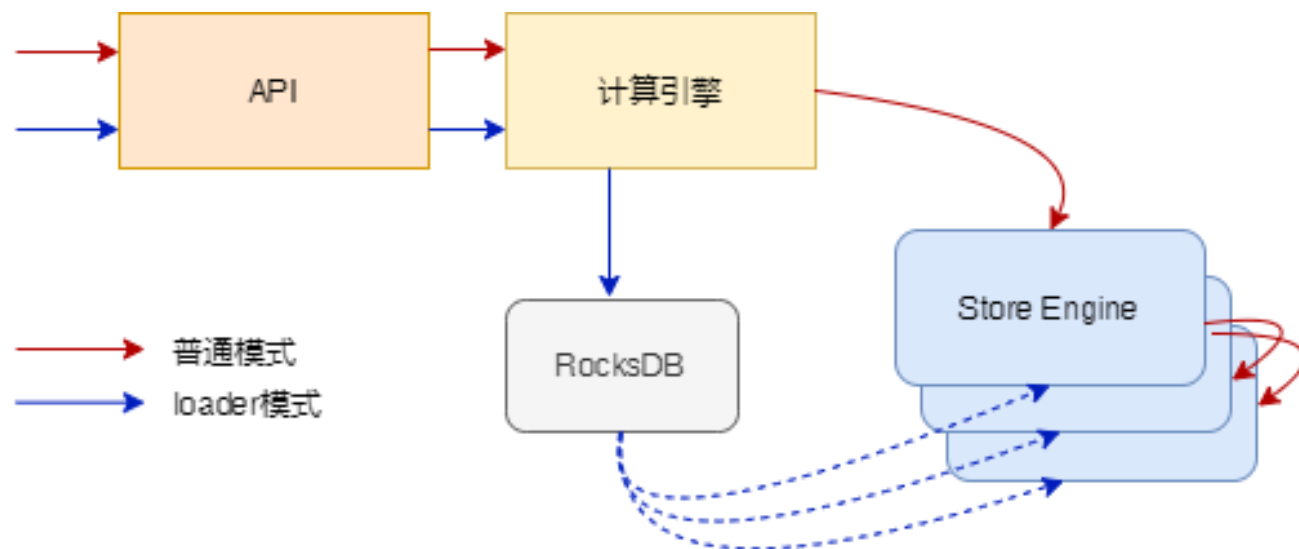


原HugeGraph中分为s、m、g三个存储端实例，CirroData-graph将实际存储顶点与关系的实例g进行开发，实现Multi-raft-group逻辑分区，保证批量插入时，每个节点都会工作。并且保证高可用，只要服务集群中超过半数节点正常，即可正常对外提供服务。使用pd集群来记录各Store中的region信息和状态，客户端访问时，通过pd获取region信息，然后向对应region的leader发起请求。

为适配CirroData-Graph，新增loader模式

普通模式下，数据经由API、计算引擎加载到一个节点的存储引擎中，然后再通过Jraft同步到其它两个节点中

因为需要保证数据的一致性，所以在数据加载方面会存在很大的性能瓶颈，为适配CirroData-Graph，我们新增了loader模式



在loader模式下，数据经由API、计算引擎先落地到RocksDB中，然后再将RocksDB的sst文件以异步任务的方式通过JRaft同步到所有节点中，这样在保证数据一致性的前提下，也让加载性能有了一个质的提高。

在loader模式下，我们的加载性能可达单机版HugeGraph性能的70%左右，这对于分布式系统来说已经达到了一个非常不错的性能。后期我们还会继续完善loader模式，争取在加载性能方面与HugeGraph持平，甚至可以在增加集群机器数量后超越HugeGraph

HugeGraph批量插入的性能

数据库	email-enron (30w edge)	amazon0601 (300w edge)	com-youtube.ungraph (300w edge)	com-lj.ungraph (3000w edge)
HugeGraph	0.629	5.711	5.243	67.033
TitanDB	10.15	108.569	150.266	1217.944
Neo4j	3.884	18.938	24.890	281.537

HugeGraph最短路径查询性能

数据库	email-enron (30w edge)	amazon0601 (300w edge)	com-youtube.ungraph (300w edge)	com-lj.ungraph (3000w edge)
HugeGraph	0.494	0.103	3.364	8.155
TitanDB	11.818	0.239	377.709	575.678
Neo4j	1.719	1.800	1.956	8.530

HugeGraph K-out性能

测试场景	维度	一度	二度	三度	四度	五度
K-out	时间	0.039s	0.045s	0.053s	1.10s	2.92s
	返回顶点数	10	140	2555	50825	1,070,230

- ❑ 批量插入性能：HugeGraph(RocksDB) > Neo4j > Titan(thrift+Cassandra)
- ❑ 遍历性能：Neo4j > HugeGraph(RocksDB) > Titan(thrift+Cassandra)
- ❑ 图常用分析方法性能：HugeGraph性能优于Neo4j和Titan，K-neighbor和K-out场景，HugeGraph能够实现在5度范围内秒级返回结果
- ❑ 社区聚类算法性能 Neo4j > HugeGraph > Titan

数据导入与核心查询功能的性能指标显著优于Neo4j。

性能比较：

数据：

点： 74314635 边： 139951301

数据：

点： 65608366 边： 1806067135

-201035082963479689,实体
-1779678833482502394,值
464640205838057683,脂肪
-1861609733419239066,别名：饴糖、糖糖、物、软糖。
-2047289935702608120,词条
5842706712819643509,词条（拼音：cí tiáo）也叫词目，是辞书学用语，指收列的词语及其释义。
-3063129772935425027,文化
-248494224944426630,红色食品
-3877061284765534378,红色食品是指食品为红色、橙红色或棕红色的食品。
-3402450096279275143,否
478618206758389997,预防感冒，缓解疲劳
-8978611301755314833,全部人群
-382812815618074210,增强表皮细胞再生和防止皮肤衰老
3455734391170888430,非科学
-4368442157131186527,生活
-401648910133347272,大龙湫
-1751058806841876591,雁荡山景区分散，东起羊角洞，西至锯板岭；南起筋竹溪，北至六坪山。
-4369745808943528904,big dragon autrum
-3278556255913778159,浙江省温州市雁荡山景区
-108136381064284954,08:00~19:00
3797530795472558959,50元
6249183780323029504,芙蓉峰

5563486759062066264,1889827357301175298,描述
5563486759062066264,5563486759062066264,名称
5563486759062066264,4942231119466354093,标签
5563486759062066264,-1966929829250449820,标签
5563486759062066264,-5659322604158396419,标签
5563486759062066264,7686044081446666806,标签
1865439276516310604,570765762641244629,描述
1865439276516310604,1865439276516310604,名称
1865439276516310604,4942231119466354093,标签
1865439276516310604,-1966929829250449820,标签
1865439276516310604,-5659322604158396419,标签
1865439276516310604,7686044081446666806,标签
7788781722555176146,-9071425129990018287,描述
7788781722555176146,4942231119466354093,标签
7788781722555176146,-1966929829250449820,标签
7788781722555176146,7686044081446666806,标签
6420045058316358689,-8255602379122600218,描述
6420045058316358689,4942231119466354093,标签
6420045058316358689,-1966929829250449820,标签
6420045058316358689,7686044081446666806,标签
-2047289935702608120,-2047289935702608120,语义关系

3200
26600
232600
1007000
61500
310000
1439000
41800
77100
233700

101 102
101 104
101 107
101 125
101 165
101 168
101 170
101 176
101 180
101 181

导入结果：

	点（vertices/s）	边（edges/s）
HugeGraph【单机】	100231	93731
CirroData-Graph【3副本】	99778	91886

	点（vertices/s）	边（edges/s）
HugeGraph【单机】	287125	204778
CirroData-Graph【3副本】	285447	199915

结论：在分布式存储3副本情况下，数据批量导入基本与原有hugeGraph单机持平。

- ✓ CirroData Graph图数据库的核心能力是为解决复杂关联关系而生的。CirroData Graph图数据模型和图数据库天生适用于以联系为中心的领域
- ✓ CirroData Graph也可以找到更多的关系，包括物与物、物与人、人与人之间的关系，这也是传统数据库无法解决的问题。
- ✓ CirroData Graph图天生没有模式。CirroData Graph数据结构能随时变更Schema。

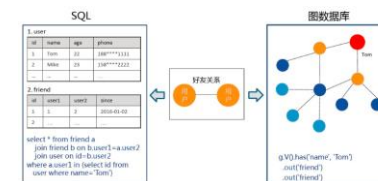
❑ **分钟级到毫秒级的性能提升：**将复杂join查询转换为图的快速遍历，而图的遍历几乎与数据集的大小关系不大。

❑ **开发周期的显著缩短：**领域专家、技术架构师和开发人员可以使用同样的模型来讨论和描述核心领域问题，减少了领域与技术之间的阻抗不匹配，没有了规范化和反规范化之间的困扰。

❑ **急速的业务响应能力：**变化的商业环境、技术以及用户行为催生系统的快速应变能力：新业务、新模式必然导致数据转移和转换、数据模型的修改，这个数据模型又与应用程序紧密相关。CirroData Graph图天生没有模式，意味着可以快速适应业务和环境的变化。

❑ **企业可用：**数据是重中之重，一个关键业务中使用的数据技术必须是健壮和可扩展的。大数据环境下的可用性不仅体现在数据上，也体现在性能响应上。

❑ **自主研发，安全可控。**CirroData-Graph是基于国人开发的HugeGraph二次开发的，开发团队深度参与，掌握自主可控、安全可信的数据库核心技术。CirroData-Graph适配多种国产操作系统以及服务器。



■ 从性能的方面考虑（涉及到数据量大小、是否为集群等因素）

- ✓ 1、Neo4j: 单机性能明显，企业版是集群模式（非分布式）
- ✓ 2、hugegraph: 单机性能（rocksdb后端）和neo4j相近甚至较好；同时还可以通过配置存储引擎，适用集群存储可支持大数据超千亿级以上（性能相对单机会有所降低）。
- ✓ 3、JanusGraph: 开源的分布式图数据库，单机性能较差。但是分布式可支持大数据超千亿级以上，和apache下的spark、hbase等结合度高
- ✓ 4、CirroDataGraph: 单机性能（rocksdb后端）和hugegraph相近甚至较好；同时还可以通过配置存储引擎，采用集群存储可支持大数据超千亿级以上（性能相对单机会有所降低）。**可以通过分布式存储达到高可用。**

■ 从功能的完备度、易用性等方面考虑

- ✓ 1、Neo4j: 功能比较齐全，但是功能都比较独立。和别的存储引擎耦合性低，不能相互组合使用。有可视化操作系统，简单功能可以实现。支持灾备，支持事务锁等。
- ✓ 2、huegraph: 功能支持健全，导入组件支持各种数据源。可视化操作组件 hubble 支持功能齐全，存储后端引擎支持宽泛，能和相关的数据库搭配使用，更适合刚上手的用户。整体上常用算法都进行封装过，易用性强。有HA组件支持灾备，事务方面的支持较弱
- ✓ 3、JanusGraph: 功能方面支持（不支持可视化界面、HA灾备）其他相关功能都具备。整体上各功能使用没有hugegraph便捷易用。
- ✓ 4、CirroDatagraph: 功能支持健全。**在hugeGraph基础上增加了分布式高可用组件（Mutl-Raft-Group）和独立的批量数据loader模式**，更加贴近现实用户场景。

图存储	Neo4j	JanusGraph	HugeGraph	CirroDataGraph
容量水平扩展	不支持	支持	支持	支持
后端存储引擎	自己独立的存储模式，自定义存储，无法使用集群存储。	支持后端存储转换(hbase、cassandra、Berkeley DB等)	支持后端存储转换(hbase、rocksdb、cassandra、postgresql、mysql、ScyllaDB)	支持后端存储转换(单机版支持hbase、rocksdb、cassandra、postgresql、mysql、ScyllaDB，分布式版暂时只支持Rocksdb)
是否支持事务	支持	支持	支持RC级别事务	支持RC级别事务
图分区	不支持	支持	支持	支持
全文检索功能	Lucene	通过ES实现	内置，易维护。	内置，易维护。
全内存式存储（效率快）	不支持	不支持	支持	支持
持久化存储	支持	支持	支持	支持
二级索引	支持	支持	支持	支持
范围索引	支持	不支持	支持	支持
高级索引（联合索引、全文索引）	支持	支持	支持	支持
图查询语言	cypher	gremlin	gremlin	gremlin
shcema管理	手动编写	支持自动创建schema，但是固定的。不建议这么做。	1、支持gremlin语言编写。2、可用可视化界面进行创建和复用	1、支持gremlin语言编写。2、可以使用可视化界面进行创建和复用。
多图实例	不支持	不支持	一个server配置多个图，可以同时应用，互不影响。	一个server配置多个图，可以同时应用，互不影响。
多ID类型（自增/主键/自定义）	不支持	不支持	支持	支持
属性类型（数字字符串日期集合）	支持	支持	支持	支持

图查询	Neo4j	JanusGraph	HugeGraph	CirroDataGraph
RESTFUL API	支持	支持	支持	支持
图处理	独立的图处理引擎	使用Tinkerpop3	使用Tinkerpop3	使用Tinkerpop3
高频图算法	通过cypher代码自行实现	通过Gremlin 代码自行实现	封装了（ShortestPath、k-out、k-neighbor等），使用更友好	封装了（ShortestPath、k-out、k-neighbor等），使用更友好
大规模查询	单机查询，数据到达一定规模，性能会降低	支持千亿级数规模查询	支持千亿级数据规模查询；支持 大规模分页查询	支持千亿级数据规模查询；支持 大规模分页查询
属性图	支持	支持	支持	支持

分布式	Neo4j	JanusGraph	HugeGraph	CirroDataGraph
分布式存储	不支持	不支持	不支持，新版支持单raft组	支持。多raft组
独立数据加载模式	无	无	无，新版增加	有
自主可控	不	不	可	可

工具链	Neo4j	JanusGraph	HugeGraph	CirroDataGraph
可视化工具	可以创建图，删除图，可以支持图数据查询。不支持数据导入等操作,不支持多图管理。	无可视化界面,但是可集成Cytoscape、Gephi等（集成操作复杂）。	提供原生的可视化界面，支持多图管理、图查询、数据导入、schema复用。易用性强。	提供原生的可视化界面，支持多图管理、图查询、数据导入、schema复用。易用性强。
高可用及容灾备份	支持灾备	不支持	支持灾备 原生HA组件	支持灾备 原生HA组件
数据导入	本地csv文件	支持本地csv等文件、hdfs导入。	支持本地csv, json, text, hdfs文件导入；也支持常用的数据库mysql、oracle、sqlServer数据导入	支持本地csv, json, text, hdfs文件导入；也支持常用的数据库mysql、oracle、sqlServer数据导入
图管理	可实现可视化界面图管理	方便从配置文件创建不同的图	1、有tools组件可以实现图的管理 2、可视化界面进行图管理	1、有tools组件可以实现图的管理 2、可视化界面进行图管理
Spark GraphX	不支持	支持	支持	支持
部署方式	部署较复杂（配置项多）	单节点部署简单，多节点部署复杂(依赖hadoop、hbase /zookeeper)	可以一键部署（修改配置可以更换存储引擎）	可以一键部署（修改配置可以更换存储引擎）
发展模式	开源版和企业版共存	开源版	有开源版和商业版共存	有开源版和商业版共存
服务支持	Neo4j	IBM	百度	东方国信



BONC 东方国信

让数据改变工作与生活



THE END