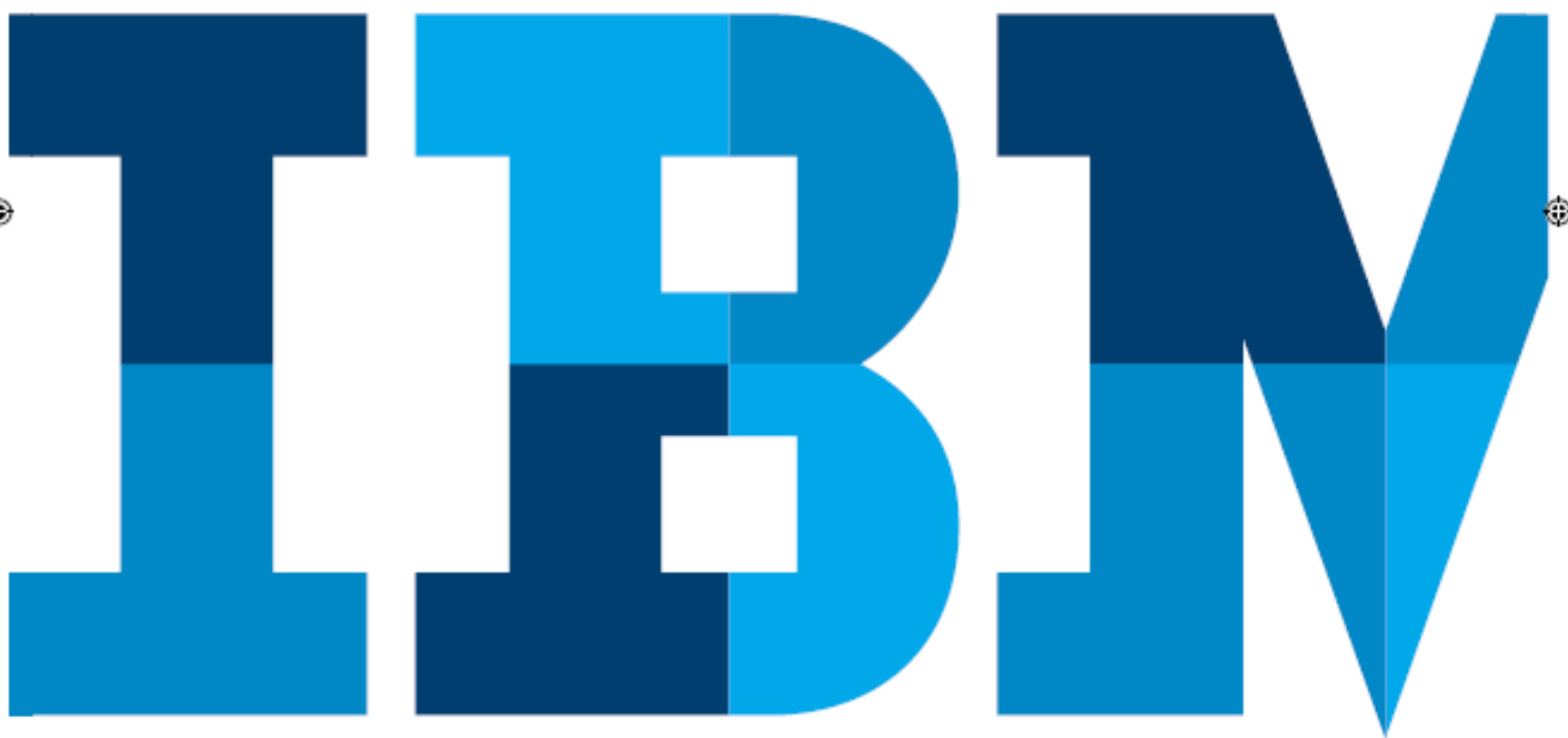


为什么选择 NoSQL ?

非关系型数据库新时代的数据库选项



目录

- 2 新型应用产生新型数据
- 2 NoSQL 发展史简介
- 3 NoSQL 数据库的类型
- 4 为什么考虑采用 NoSQL ?
- 5 总结
- 5 IBM Cloudant 入门介绍
- 5 有关更多信息

新型应用产生新型数据

如今，新的网络应用、移动应用和物联网 (IoT) 应用层出不穷；同时，新兴的趋势也在改变网上消费者的行为，并孕育出新类型的数据。这些因素促使开发人员不得不重新评估数据的存储和管理方式。为了满足当今的需求，企业的数据库必须能够提供一个灵活、可扩展的解决方案，来高效安全地管理流入和流出全球用户群的大规模数据流。

开发人员和 IT 人员都发现很难在快速将所有这些数据融入关系模型的同时，动态地扩展数据库以满足用户的性能要求。某些情况下，这甚至是不可能完成的任务。于是，许多人瞄准了 NoSQL 数据库的灵活性，这也是为什么有机构预计到 2020 年 NoSQL 市场价值将翻一番、达到 34 亿美元的一个重要原因。

NoSQL 发展史简介

NoSQL 亦称为“不只是 SQL”或非关系型数据库。网站的兴起带来了激增的数据类型、数据访问方式和数据可用性。为了处理这些问题，NoSQL 应运而生。

20 多年前，如果应用架构师和开发人员需要为其应用寻找一个数据存储数据库，他们会从五花八门的关系型数据库中选择一款。事实上，自二十世纪 70 年代以来，关系型数据库就一直是他们的实际选择，因为这也是开发人员（在四处教授计算机科学课程）和基础架构团队（非常了解工具和可预测的运行特性）的唯一选择。Oracle、MySQL、SQL Server 和 IBM® DB2® 就是一些最热门的关系型数据库。

但是，到了二十世纪 90 年代末和二十一世纪初，互联网应用和互联网企业迎来了爆炸式的发展，各种应用层出不穷，有服务企业内数千名内部员工的应用，也有服务公共网络上数百万用户的应用。对于这些应用来说，性能和可用性排在首位。为了解决新兴的海量数据的高可用性问题的，谷歌、Facebook 和亚马逊等企业不得不开发新技术。值得庆幸的是，他们记录下了自己的工作、发布了白皮书，并公开了他们的技术，让整个互联网行业都能够基于他们的成果继续开发。到二十一世纪前十年的最后几年，市场上已经有了几种新的非关系型数据库技术，而 NoSQL 就是所有这些技术的一个总称。

NoSQL 的根源：开源技术

许多 NoSQL 数据库都起源于开源社区。正是这种传统使得 NoSQL 数据库越来越普及，采用率也越来越高。我们经常看到一些企业推出了一个附带 NoSQL 服务和支持的商业版数据库，但同时又加入了其开源合作伙伴的社区，比如，Datastax for Apache Cassandra 和 IBM Cloudant® for Apache CouchDB，甚至 MongoDB 都有一个开源版本的 MongoDB 软件。

NoSQL 数据库的类型

NoSQL 只是一个用来描述所有非关系型数据库的术语。NoSQL 中的技术、数据类型和使用案例又有很大区别，因此人们对 NoSQL 数据库进行了分类。公认 NoSQL 数据库有以下四大类：

键值存储数据库

这类数据库为每一个键赋予了一个值，比如计算机编程技术中的散列表。为了更好地了解这类数据库，我们可以用文件系统来进行类比。路径就好比键，内容就相当于文件。通常，我们不需要更新字段。如果需要进行变更，我们需要更新的也是整个值，而非键。这大大地简化了数据库的扩展。但是，同时，它也限制了键值存储数据库的查询功能和其他高级功能的复杂性。单纯的键值存储数据库包括 MemcacheD、REDIS 和 Riak。

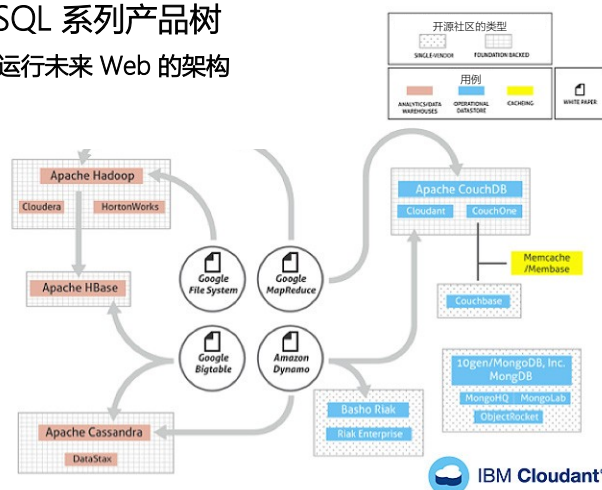
图形存储数据库

图形数据存储数据库非常擅于处理互连的数据。图形数据库提供了各个节点之间的连接（在图形数据库中称为“边”）。节点和他们的边都能存储其他属性，比如键值对。图形数据库的优点是能够沿着节点之间的连接进行传输。但是，它一般需要所有数据匹配一台机器，这限制了它的可扩展性。Neo4j 和 Sesame 就是典型的图形存储数据库。

列存储数据库

关系型数据库将所有数据存储于磁盘的某个表行中，这样您就可以快速地进行行检索。列式数据库通常会将某一个行的所有值序列化地存储在磁盘上，让您能够根据特定属性快速检索聚合数据。在数据仓储和分析场景中，您可能会针对某个值进行范围查询，因此，前述功能在这些场景中非常有价值。比如，Hbase 和 Cassandra 就是这样的列存储数据库。

NoSQL 系列产品树 了解运行未来 Web 的架构



文档存储数据库

文档数据库以文档形式存储记录。通常，您可以将文档视为一组键值对。键一般以字符串形式存在，值可以是字符串、数字、Boolean 类型、阵列和其他嵌套的键值对。值可以嵌套至任意深度。文档数据库常用的语法包括 XML 和 JavaScript Object Notation (JSON)。MongoDB、CouchDB、Cloudant 和 MarkLogic 都是典型的文档数据库。

为什么考虑采用 NoSQL ?

尽管现有的 NoSQL 数据库之间差别很大，但这类数据库都具有以下特性：灵活性、可扩展性、可用性、更低的成本和特定功能。

灵活性

灵活的模式和直观的数据结构是吸引开发人员在敏捷开发周期中开发软件的重要特性，而大多数 NoSQL 数据库都具备这些特性。与关系型数据库不同，NoSQL 数据库在更新模式时一般不会引起停机。得益于这种灵活性，NoSQL 能够在变更代码库时，消除复杂的数据库模式变更流程。因此，它常常被用于支持敏捷开发实践。

可扩展性

这里的扩展不只是指扩展数据规模，还包括扩展使用这些数据的并发用户。相比关系型数据库，通常 NoSQL 数据库能为涉及扩展的各种使用场景提供更专业的支持，同时也能以更简单的方式开发相关的应用功能。

许多 NoSQL 数据库都能够实现水平扩展。在将数据分散到服务器集群中时，它们采用的方式也比关系型数据库更简单。关系型数据库在跨片区 (shard) 执行 JOIN 查询时，性能会有所降低。但是 NoSQL 数据库能够完全避免 JOIN，保持高性能状态。

可用性

停机会导致收入和客户流失，用户也会感觉沮丧不已。幸运的是，一些 NoSQL 数据库提供了新的或不同的复制架构，提高了不同类型的 NoSQL 数据库在读取之外的写入方面的可用性。也就是说，当一个或多个数据库服务器/节点出现故障时，系统中的其他节点能够继续运行，而不损失数据。

更低的运营成本

因为 NoSQL 根植于开源技术，因此，其入口点成了吸引用户一大亮点。我们经常听到客户采用 NoSQL 后，成本比当前的数据库低了很多，同时还能收到相同甚至更卓越的性能和功能。过去，大型的关系型数据库都是基于昂贵的机器和大型机运行。而 NoSQL 数据库的分布特性意味着，它们的部署和运营可以在云端的服务器集群上进行。

NoSQL JSON 文档存储数据库的亮点

- JSON 模式可以在无介入的情况下快速发展
- 它只有 6 种值，因此实施和使用都很方便
- JSON 具有自描述性，非常容易理解
- 它有助于数据库性能和可扩展性的提升

示例：NoSQL JSON 文档存储数据库与关系型数据库管理系统 (RDBMS)

RDBMS 方法是将关系分解至几个不同的表中（在不同的逻辑表中存储不同的记录，以提高一致性），而 NoSQL 文档存储数据库则采用了一种非常规方法，它会将记录存储在一个个单独的 JSON 文档中（将记录及其所有相关数据存在一起）。

这样，借助 NoSQL，您可以一次性获取和导入所有对象，无需从不同的表格中汇总数据。对于数据库集群来说，关系型数据库的 JOIN 无疑是影响可扩展性的最大绊脚石。

特定功能

NoSQL 数据库各有优势。为了刺激消费和帮助用户实现更紧密的集成，许多 NoSQL 数据库提供商都在数据库中加入不同的专有功能。比如，针对地理空间数据的专有索引和查询功能；针对搜索的集成式全文索引功能；自动化数据复制和同步功能；以及应用友好型 RESTful 网络应用编程接口。

总结

如今，您为下一个应用选择的数据库变得空前的重要。幸运的是，您拥有不止一种选择。那是因为如今的应用不仅要能持续运行，还必须能够高效管理持续增加的海量多结构数据。这种情况下，NoSQL 不再只是一个热门词汇。不论是小商店还是大企业，他们在采用数据库时，都会认真考虑 NoSQL。

尽管 NoSQL 数据库有一个共性，比如都属于非关系型数据库，通常都便于扩展；但是它们也有一些独一无二的差异之处。因此，您在选择合适的 NoSQL 解决方案时，要仔细考察这些差异之处，才能满足您独特的数据需求。理想的 NoSQL 数据库可能是替代关系型数据库的可行方案，也可能辅助现有系统，与现有系统同时运行。

高效交付数据是大势所趋，如果您的应用开发项目涉及扩展、不同的数据类型或者大量潜在用户，那么您在计划这些项目时，请务必考虑 NoSQL 技术。

IBM Cloudant 入门介绍

IBM Cloudant 是一款完全托管的 NoSQL 数据库即服务 (DBaaS) 产品，提供快速的全套数据管理服务，让您高枕无忧。同时，它还提供 Cloudant Local 版本，该版本将 Cloudant 平台的功能全部部署在您的私有数据中心内。甚至您可将 Cloudant Local 和 Cloudant Managed DBaaS 数据库连为一体，形成混合云数据库，进而实现云成本、覆盖范围、性能和合规控制之间的最佳平衡。立即访问 cloudant.com，注册获取免费账号并开始使用。

有关更多信息

有关更多信息，请联系您的 IBM 销售代表或 IBM 业务合作伙伴，或访问以下网站：cloudant.com 或 ibm.com/cloudant。



© Copyright IBM Corporation 2015

IBM Corporation
Software Group (或相应事业部或无事业部)
Route 100
Somers, NY 10589

美国印刷
2015 年 2 月

IBM、IBM 徽标及 ibm.com 是 International Business Machines Corporation 在世界各地司法辖区的注册商标。Apache Software Foundation、“Apache Cassandra”、“CouchDB”及“Apache CouchDB”是 The Apache Software Foundation 的商标或注册商标。IOS 是 Cisco Systems, Inc. 在美国和其他国家/地区的商标或注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。Web 站点 www.ibm.com/legal/copytrade.shtml 上的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表。

本文档截至最初公布日期为最新版本，IBM 可随时对其进行修改。IBM 并不一定在开展业务的所有国家或地区提供所有这些产品或服务。

本文档内的信息“按现状”提供，不附有任何种类的（无论是明示的还是默示的）保证，包括不附有关于适销性、适用于某种特定用途的任何保证以及非侵权的任何保证或条件。IBM 产品根据其提供时所依据协议的条款和条件获得保证。

1 www.marketresearchmedia.com/?p=568



请回收利用