



## 要点

- 单独或共同对多个数据库和大型数据存储库进行查询
- 集中实施访问控制与监管
- 对于应用而言，众多数据库（即使是全球分布的数据库）就像是单一数据库
- 使用易于扩展、功能强大的平台，简化数据分析

## 背景

数据无处不在，目前世界上最优秀的企业都以数据作为主要推动力。企业从越来越多、越来越多样化的数据源收集数据，分析和开展业务运营，所涉及到的数据源数以千计，甚至数以百万计。集中收集、管理、存储、处理和分析数据的复杂性、成本、时间和错误风险呈指数级上升。与此同时，作为数据来源的数据库和存储库的功能也越来越强大，它们本身就具有充足的数据处理和存储能力。

## 数据虚拟化概述

IBM® Cloud™ Private for Data 中的数据虚拟化功能名为 IBM Data Virtualization，这是一种独特的新技术，它将所有这些数据源连接到一组能够实现自我平衡的数据源或数据库（称为群集）中。请参阅图 1。不需要将数据复制和存储到集中位置再进行分析查询。分析应用提交查询，在数据源所在的服务器上处理分析工作。查询结果合并到群集中，并返回给原始应用。因此不需要复制任何数据，数据始终保留在数据源中。

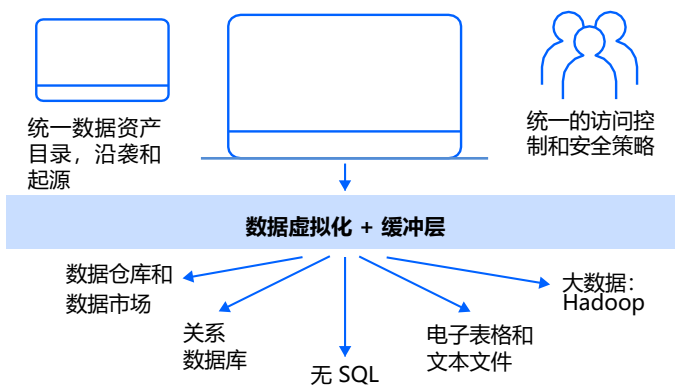


图 1. IBM Cloud Private for Data 中的数据虚拟化

## 数据虚拟化之工作原理

应用连接到 IBM Data Virtualization 时，就像是连接到单一 IBM Db2® 数据库那样简单。进行连接后，应用可以提交针对系统的查询，就像查询单一数据源数据库那样。具有与查询相关的数据的所有参与数据源将以协作方式分配和计算工作负载。

## 重要功能

IBM Data Virtualization 包含许多重要功能，旨在帮助企业更高效地处理数据。

## 协作式计算

通过使用每个数据源的处理能力，在每个数据源的实际存储位置访问数据，可以避免移动和复制数据所造成的延迟。此外，所有存储库数据都可实时访问，因此基本上消除了监管问题以及数据出错问题。由于不需要抽取、转换和加载 (ETL) 以及复制数据存储，因此能够加快处理速度。与现有方法相比，该流程能够更迅速、更可靠地为决策应用或分析人员提供实时洞察。它还与现有方法高度互补，因为有时候出于历史记录、归档或监管的目的，还是有必要复制和移动一些数据的。

## 模式折叠

在分布式数据系统中，众多数据库采用相同模式存储数据是很常见的。例如，您可能有多个存储销售数据或交易数据的数据库，每个数据库存储一组租户或一个地区的数据。IBM Data Virtualization 可以自动检测各个系统中的相同模式，并使它们在数据虚拟化过程中以单一模式出现 — 这个过程就称为模式折叠。例如，如果 20 个数据库中都存在的 SALES 表，现在可以作为单个 SALES 表出现，并可通过“结构化查询语言” (SQL) 作为一个虚拟表进行查询。

## 简单的连接视图工具

精巧的内联工具支持针对不同类型而且可能分布在不同地理位置的数据库定义表视图（如图 2 所示）。

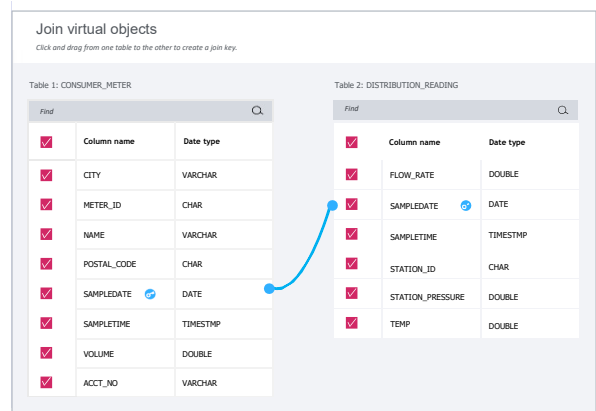


图 2: 直观的界面使连接表视图更为简单



IBM Cloud PGQ for Data 中的 IBM Data Virtualization 支持一系列主要目标，包括：

- 实现现代化，更快速、更轻松地将交付现代互动系统
- 实现实时分析，满足企业的即时需求
- 实现优化，降低访问组织数据的成本和复杂性

**IBM Data Virtualization 有助于实现自助式 BI。**可复用的虚拟数据资产以适合业务部门的形式呈现数据，使用户能够轻松与数据交互，而无需了解物理数据层的复杂性或数据存储的位置。它还支持多个 BI 和报告工具从数据虚拟化层获取数据。

**IBM Data Virtualization 提供了统一的全方位视图。**虚拟化的数据资产能够实时提供完整的数据视图。虚拟数据层作为业务信息的统一集成视图，可以帮助用户更好地理解和使用组织数据。

**IBM Data Virtualization 提供敏捷的面向服务的架构 (SOA) 数据服务。**数据虚拟化层为 SOA 应用提供数据服务层。它有助于加快虚拟资产的创建，不需要触及底层数据源，能够自动发现和映射封装了数据访问逻辑的元数据。数据虚拟化还支持多个业务服务从集中位置获取数据，并在业务服务和物理数据源之间建立松散耦合。

**IBM Data Virtualization 改进了信息控制。**它通过提供集中式访问控制，建立强大的安全基础架构以及减少数据的物理副本来提高数据质量，降低风险。元数据存储库对组织的数据存储以及各种数据存储中数据之间的关系进行编目，从而实现透明度和可视性。

### 总结：转变决策模式，加快决策速度

IBM Cloud PGQ for Data 中的数据虚拟化技术是帮助企业实现以下目标的理想之选：

- 实现盈利，促进增长和风险降低
- 提升敏捷性和生产力
- 优化现有 IT 投资

SQL (API)

IBM Data Virtualization

<https://www.ibm.com/cn-zh/analytics/data-virtualization>

免费咨询电话：400-669-0260



© Copyright IBM Corporation 2019

IBM Corporation  
New Orchard Road  
Armonk, NY 10504

美国出品  
2019 年 1 月

IBM、IBM 徽标、ibm.com、Cognos、Db2、IBM Cloud、IBM Watson 和 Informix 是 International Business Machines Corp 在全球许多司法管辖区的注册商标。其他产品和服务名可能是 IBM 或其他公司的商标。Web 地址 [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml) 的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表。

Netezza 是 IBM International Group B.V., an IBM Company 的注册商标。

Microsoft、Excel 和 SQL Server 是 Microsoft Corporation 在美国和/或其他国家或地区的商标。

本文档为自最初公布日期起的最新版本，IBM 可随时对其进行更改。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。

评估和验证任何其他产品或程序与 IBM 产品和程序一起运行的情况，则由用户自行负责。本文档内的信息“按现状”提供，不附有任何种类（无论是明示的还是默示的）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。

良好安全实践声明：IT 系统安全性涉及通过预防、检测和响应来自企业内外的不当访问，保护系统和信息的安全。不当访问可能导致信息被篡改、损毁或滥用，也可能导致您的系统遭到损坏或误用，包括用于攻击他人。没有任何 IT 系统或产品可视为是完全安全的，也没有任何单一产品、服务或安全措施能够完全有效地预防不当访问或使用。IBM 系统、产品和服务旨在成为全面的合法安全方法的组成部分，这必然会涉及额外的操作程序，还可能需要其他系统、产品或服务才能达到最佳效果。IBM 不保证任何系统、产品或服务免受任何一方恶意或非法行为的影响，或使您的企业免受任何一方恶意或非法行为的影响。

关于 IBM 未来方向和意向的声明仅表示目标和目的，可能随时更改或撤销，恕不另行通知。