



主要特点

- 基于 IoT、Web 和移动应用的参与型系统生成的数据应在其原生环境，即云中进行管理和维护，而不是移动到本地部署的系统进行分析。
- 在云中保留数据可实现按需分析，并可利用 JSON 数据库服务和关系云数据仓库之间的简单集成，同时只需移动或转化最少量的数据。
- 将数据库管理交给专业的服务提供商，您就有时间和精力将注意力重新放在可带来市场竞争优势的活动中。

分析 JSON 数据以获得全方位客户洞察

将分析引入物联网、Web 和移动数据以促进业务增长

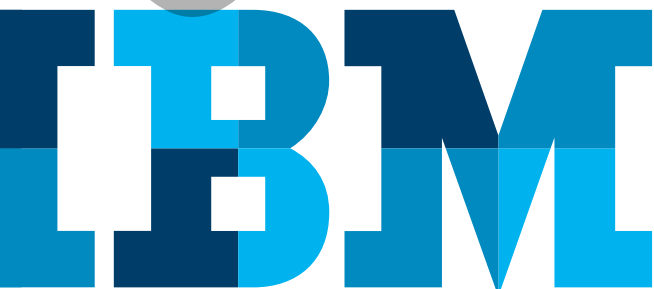
概览

当今的物联网 (IoT)、Web 和移动应用程序可收集比以往更多的客户相关数据。这种庞大的信息量使我们有比以往更加细致地研究用户及他们的需求和习惯。

然而，这种数据的庞大容量、速度和多样性必然使其难以进行分析。那么，如何高效、安全地存储半结构化 NoSQL 数据表单，例如来自 Web 和移动应用程序的 Java Script 对象符号 (JSON) 文档 - 更进一步说，如何生成可促进业务增长的有意义的客户洞察？

有许多影响因素造成了这种挑战：

- 参与型系统创建半结构化数据（来自 IoT、Web 和移动数据源）的大型存储库，这对传统的数据管理造成了阻碍。规划、构建和维护此类数据的基础架构会占用大量资源，而这些资源本来更适合用于吸引客户。
- 参与型系统生成的数据往往宽度和深度均有所不足，且存储在简化的架构中，这种架构缺少应用扩展模型和预测算法所需的稳健性。参与型系统和记录型系统（例如客户关系管理 (CRM)、主数据管理 (MDM)、库存和其他事务性系统）之间缺乏兼容性给分析工作带来了阻碍，而这种分析工作本来有助于全方位地了解客户。
- 本地部署的传统记录型系统往往以可预测的稳定速率发展和变化，而参与型系统产生的数据更具流动性。随着新技术的出现并产生半结构化的数据，存储此类数据以及从中得到有用的客户洞察变得更加困难。



幸运的是，出于以下三点关键原因，最近可组合云服务的爆炸式增长提供了一些应对 NoSQL 分析挑战的极具吸引力的方式：

- 基于 IoT、Web 和移动应用的参与型系统生成的数据应在其原生环境，即云中进行管理和维护，而不是移动到本地部署的系统进行分析。
- 参与型系统生成的数据缺乏传统记录型系统所具备的成熟架构，这样就很难使用关系数据仓库存储此类数据以便进行分析。但是，理想的处理方式是，在同时利用 SQL 和 NoSQL 数据库系统的原生云解决方案中存储此类数据。
- 借助正确的平台，由全新的颠覆性技术创建的未来数据可在基于云的数据仓库中得到充分利用。

数据孤岛：影响 NoSQL 分析的障碍

Web 或移动应用程序产生的数据存储可调整规模和不断增长。通过规划、构建和维护可扩展的基础架构引导这种大规模涌入的数据需要占用相当庞大的资源。例如，因为您的服务器还无法支持额外的数据流量，所以您不得不避开使用新的 IoT 技术，又或者是延迟启动 Web 或移动应用程序。

那么，一旦收集了客户相关数据，如何从中得到有意义的洞察？来自 IoT、Web 和移动数据源的许多数据都具备半结构化特性。例如，JSON 数据存储于架构不断改变的庞大存储库中，其缺乏大多数记录型系统（如 CRM、

库存或财务系统）中数据的关系结构。因此，很难通过商业智能 (BI) 和分析工具探索 JSON 数据，因为这些工具往往依赖于传统的架构。

此外，从 IoT、Web 和移动应用程序收集的数据以客户的在线互动和行为作为基础，因此只能提供一定程度上片面的客户洞察。为得到更加深入、更具意义的客户洞察，需要将来自参与型系统的客户数据与来自记录型系统的数据进行比较和对照，形成关于未来客户需求和行为的更细致入微的洞察。

为消除参与型系统与记录型系统之间的鸿沟，我们需要应对许多挑战。两种系统在如下领域存在差异：

- **文化**。来自参与型系统的数据属于 Web 开发人员，而记录型系统则是传统 IT 人员的领域。
- **实践**。很难将 NoSQL 数据库中采用灵活架构的半结构化数据合并到 SQL 数据库中采用严格架构的高度结构化数据中。
- **数据传输**。支持参与型系统的服务器可能位于任何位置，而来自记录型系统的数据通常存储在受防火墙保护的安全位置中，例如本地部署的数据中心或私有云。

最终，大多数公司都形成了数据孤岛，其结合来自各种参与型系统和记录型系统的信息并进行分析的能力相当有限。但是，如果可以提取 JSON 文档中的半结构化数据，梳理其结构，然后将其轻松导出至基于云的数据仓库中，情况又会如何？这样的话，就可以在原生环境，即云中分析数据，甚至是结合来自后端系统的数据进行统一分析。将分析引入 JSON 数据可大幅度提升对客户的全局了解。

特色服务：IBM Cloudant



IBM Cloudant 是完全托管的 JSON 文档储存库，其针对处理大量的并发读写操作进行了优化，在使用大型的全球移动应用程序时通常会产生这种工作负载。

Cloudant 的作用相当于 NoSQL 数据库即服务 (DBaaS)，可提供快速数据配置和无忧数据管理。IBM 也提供了 Cloudant Local，该服务将 Cloudant 平台的强大功能引入客户拥有的数据中心内部。您甚至可将 Cloudant Local 和 Cloudant DBaaS 链接在一起，形成混合云系统，从而实现最出色的成本、范围和性能平衡以及合规控制。

灵活的 NoSQL 数据层

Cloudant 拥有在 Apache® CouchDB 基础上构建的稳健 RESTful API，用户可轻松使用任意语言或平台即服务 (PaaS) 对其进行访问。该服务包含实现高级地理空间索引的集成 API，并具备基于 Apache® Lucene 的全文本搜索功能。Cloudant 独有的多主机复制和地理负载平衡功能使其可在全球的数据中心和设备上分布数据，同时可持续扩大规模以及不间断提供数据（无论用户是否在线）。

得益于其支持的 SLA、高可用性和正常运行时间，Cloudant 提供高于行业标准的服务水平，IBM 专家可全天候管理 NoSQL 数据层，因此客户无需在此方面投入时间和精力。

极富弹性的全球可扩展性

Cloudant 避免了大量的基础架构前期资本开支，它提供按容量增长付费方式，您可根据自身需求增删计算资源。除了在数据量方面具有弹性之外，Cloudant 还可灵活地在多台服务器上分布数据，或者在多个数据中心上分布数据，使数据更接近您的用户。该服务甚至可将数据限制在特定的地理位置，从而遵从某个国家特有的数据法规。

第一步：将数据存放在云中

为 NoSQL 数据构建分析解决方案，首先需要重点考虑的是在何处存放数据。一些组织仍然致力于构建服务器农场并在企业内部存放云端生成的数据，但这种方式已被证明代价昂贵、难以处理且耗费资源。更好的方式是在云中存放云端生成的数据。

云计算现已商品化，许多大型和小型云服务提供商均提供此项业务。然而，商品化并不能排除在选择云服务提供商时进行尽职调查和规划的需要。关于选择哪个提供商作为合作伙伴的决策应符合如下四点要求：

- **服务水平。** 确立您将需要的服务水平（特别是正常运行时间和网络可靠性）；然后，将它们与云服务提供商保证遵守的服务水平协议 (SLA) 进行比较和对照。
- **可扩展性。** 随着业务的增长，您的云提供商必须能够与您同等规模地成长。确认服务器具有满足业务增长需求的可用空间，并且确认根据需要调整服务器规模所带来的额外财务成本。
- **安全。** 验证云服务具备适当的安全措施以保护敏感的客户数据。云服务如何管理数据加密、复制和灾难恢复？这些安全措施是否满足您的需求？
- **支持。** 一些云服务提供全套解决方案，而其他服务则提供多层支持结构，随时有专家响应客户的升级需求和其他特殊需求。进行了解，确定潜在的云合作伙伴可提供您需要的支持水平和规模。

经常会被忽略的一个成本领域是数据传输和网络费。许多公司错误地将 IoT、Web 和移动服务生成的数据视为临时存储在云上的内容。他们认为，为使这类数据发挥作用，需要将其导出至本地存储库，从而产生了传输该数据的相关成本。我们下面会看到，这种观念正在发生改变。

第二步：简化 JSON 数据的分析

如前所述，当半结构化数据在 NoSQL 数据库中存储为 JSON 文档时，许多成熟的 BI 工具很难对其进行分析。以前，对此类数据进行分析需要采用多点解决方案，以便从 NoSQL 数据库中导出数据，对相关数据进行转换和整理，然后将其存入本地部署的数据仓库。然而，全新的混合解决方案集成了 SQL 和 NoSQL 功能，可将基于云的数据保留在云端以便进行按需分析，从而只需移动或转化最少量的数据。

这种解决方案利用了 JSON 数据库服务 (IBM Cloudant) – 充当运营数据存储 (ODS) – 和关系云数据仓库 (IBM dashDB) 之间的简单集成，为半结构化数据提供结构，以便进行分析。此解决方案还利用了 Cloudant 的自动化架构发现过

程 (SDP)，该过程可扫描 JSON 数据库并感知数据中隐含的结构。然后，该过程在关系数据仓库中创建对应的架构，并以 BI 和分析工具可轻松理解的格式复制数据（参见图 1）。

这种方式有许多优点，包括：

- 轻松可视化 IoT、Web 和移动系统产生的最新数据并提供相关报告。
- 即时数据分析 – 基于云的数据集可能非常庞大、无组织和快速变化，因此需要利用一组同等快速的 BI 工具及时确认发展趋势和相关商机。
- 消除数据传输和网络成本。

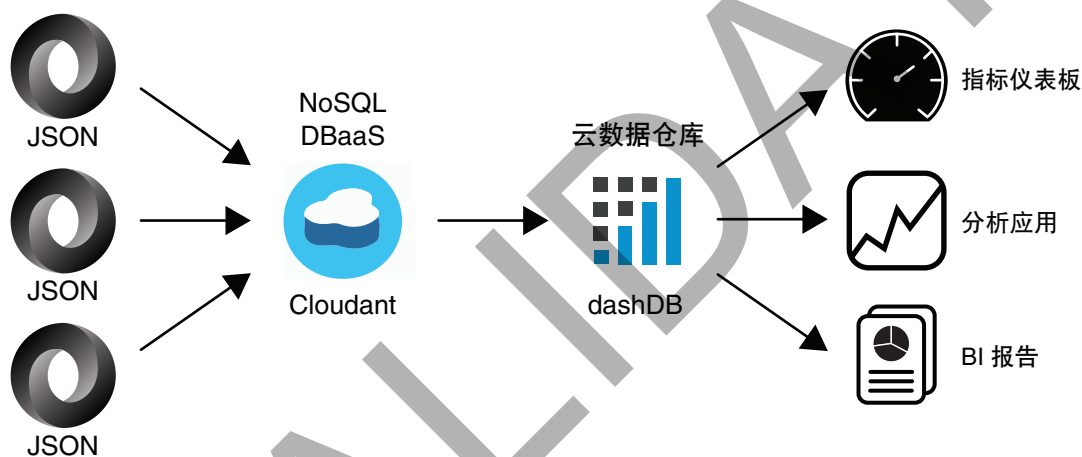


图 1：Cloudant 和 dashDB 的集成实现了对 JSON 文档数据的分析

IoT 用例：实时零售分析

最快速增长的 IoT 实施方式之一是零售行业对低功耗蓝牙 (BLE) 信标的使用。这些信标是不引人注目的连接设备，可安装在环境内的间隔处，并以单向或双向方式将信号广播到带数据连接的接收移动设备。

零售商发现 BLE 信标可高效地追踪客户在店铺内的行为。一些新技术会产生最适合存储在云中的大量数据，而 BLE 信标就是此类技术的绝佳示例。与此同时，这些信标也证明了为何应将数据保留在云中以供可行性分析使用。

例如，靠近折扣货架的信标可连接客户移动设备上的应用，即时向客户回馈所销售商品的额外电子折扣券。这种互动可作为半结构化 JSON 文档存储在 NoSQL 数据存储（例如 Cloudant）中。随着客户在一个或另一个信标的范围内移动，零售商可获得关于客户店铺内行为的一定程度的智能分析，而客户可获得高度量身定制的店铺内购物体验。

但是，只有将信标和移动应用程序生成的半结构化数据与来自自身记录型系统的结构化数据相结合，零售商才能获得真正的客户洞察。在此案例中，零售商可将来自主数据库 (MDM) 数据库的客户档案放入云数据仓库，例如

特色服务：IBM dashDB



IBM dashDB 是完全托管的云数据仓库服务，其具备如下特点：

- 快速。可通过单击一次按钮在几分钟内完成部署。
- 可扩展。按需付费，无资本支出，也不需要规划。
- 灵敏。基于 IBM 的内存分析技术不间断载入数据并立即执行，同时具备与设备类似的简便性。
- 灵活。完全兼容各种 IBM BI 和分析工具，包括 Watson Analytics 和 Cognos，以及诸如 Tableau 和 Looker 这样的第三方解决方案。

dashDB 是原生兼容 Cloudant，从而可轻松地将 Cloudant JSON 文档引入云中的关系数据仓库。这样便可根据 Web 和移动数据即时创建模型和执行极富洞察力的分析，从而可专注于促进业务增长，而不是构建新的基础架构。

dashDB 最具价值的功能之一是其能够通过单击一次按钮来比较、对比和查询多个来源的数据。dashDB 不仅兼容 Cloudant 以 JSON 文档格式存储的数据，还兼容来自本地部署的传统记录型系统的数据。

dashDB。个人资料可显示客户的身份，过去购买的衣服（包括款式、颜色和尺码），以及此次客户进店是否可带来商品的成功售出。信标产生的数据和详述客户最近拜访零售商店行为的移动数据在 Cloudant 中存储为 JSON 文档，这些数据也可通过 Cloudant 的 SDP 功能（前面已介绍过）移动到 dashDB 中。参与型系统和记录型系统中的数据以前以孤岛的形式存放，而现在则有效地结合在一起。

现在可使用 BI 工具结合和分析数据，有效地生成新的全方位客户洞察：客户是否试穿过某件衣服？他们是否购买了这件衣服？如果没有购买，原因何在？商店是否有客户所需尺码的衣服？一张极富洞察力的全方位客户视图正在逐渐成形（参见图 2）。

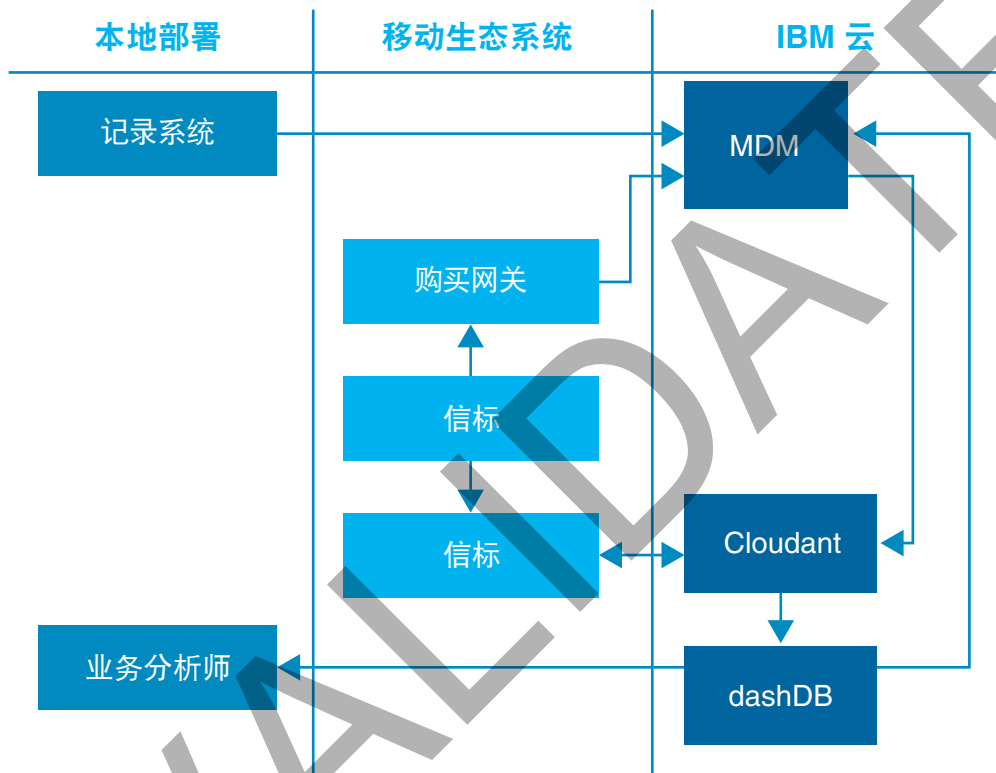


图 2：使用 IBM 云数据服务的一个 IoT 分析基础架构示例

更充分地利用应用程序数据

随着 IoT、Web 和移动应用程序以及它们所产生的半结构化数据的不断增长，基于云的数据管理的重要性更加凸显。云服务提供商可更高效地管理以较新格式（例如 JSON）存储的数据，这是他们的核心竞争力所在。将云存储管理交给专业的提供商，您就可以将注意力重新放在可带来市场竞争优势的活动中。

传统的观念将云数据存储直接视为 Web 和移动数据的存放位置，只有将其下载到本地系统，才能最充分地利用来自 IoT、Web 和移动应用程序的数据；然而，您必须接纳如下观点：使数据保持在线，并将分析引入当前存储数据的位置，即云端。采用这种方式，即可有效地将半结构化数据（例如 JSON 文档）整合到业务流程中，并更为快速地转换为可执行的洞察力。

IBM 云数据服务简介

IBM 云数据服务可为开发人员、数据科学专家和分析架构师提供一组广泛的、可组合的集成数据服务，这些服务涵盖了内容、数据和分析。我们的开放、自助式云数据服务可加快产品投入市场的速度，延长正常运行时间，提供更高的价值，并由 IBM 技术领导者提供可靠的支持。如需更多地了解 IBM 云数据服务可如何改变为开发人员创建和提供服务的方式，请关注我们的 Twitter: @IBMdashDB 和 @IBMcloudant，并访问 ibm.biz/clouddataservices。

了解更多信息

如需了解有关我们提供此服务的更多信息，请联系您的 IBM 销售代表，或访问以下网站：ibm.com/gbs/contact



© IBM 公司版权所有 2016

IBM Corporation
Software Group (或相关部门, 或无部门)
Route 100
Somers, NY 10589

美国印制
2016 年 5 月

IBM、IBM 徽标以及 ibm.com 是国际商业机器公司的商标, 已在全世界许多管辖地区注册。其他产品和服务名称可能是 IBM 或其他公司的商标。当前的 IBM 商标列表请见网站的“版权和商标信息”版块:
www.ibm.com/legal/copytrade.shtml。

本文档包含截至发布之日的最新信息, IBM 可能随时更改。本文档所述各种产品或服务不一定在 IBM 开展业务的所有国家/地区均有提供。

本文所载信息按“原样”提供, 不做任何明示或暗示的担保, 包括对适销性、特定目的的适用性的任何担保, 以及针对非侵权的任何担保或条件。IBM 根据产品交付协议中规定的条款和条件为产品提供担保。



请回收再利用
