

技术概览：IBM Cloudant DBaaS 剖析

Cloudant DBaaS 如何实现卓越可扩展性、高可用性和性能



目录：

- 1 管理数字体验时代的数据
 - 1 Cloudant 的 DNA
 - 2 Cloudant 性能特点
 - 2 Cloudant Data Layer 工作原理
 - 3 创建 Cloudant Managed DBaaS 帐户
 - 3 存储 JSON 数据
 - 4 建立数据索引
 - 5 通过 RESTful API 访问 JSON 数据
 - 5 复制数据
 - 5 监控和可视化
 - 6 扩展 Cloudant
 - 6 部署选项
 - 6 私有云要求
 - 7 Cloudant 入门
-

管理数字体验时代的数据

如今的应用程序需要管理各种结构化和非结构化数据，由用户、设备和经营场所甚至是传感器、车辆和互联网商品构成的庞大网络对这些数据进行访问。

IBM® Cloudant® 不仅仅是个数据库。它是您的应用程序的完整数据层，可作为全管理型数据库即服务 (Cloudant Managed) 提供，也可作为您自己安装并运行的 IBM® Cloudant® Data Layer Local Edition (Cloudant Local) 系统提供。Cloudant 支持弹性扩展和轻松构建文本搜索和地理位置服务等高级功能，同时可根据用户需要随时随地不间断地向设备提供数据，有助于开发者为用户打造更丰富的数字体验。

本白皮书阐述了 Cloudant Data Layer 的含义、特别之处和使用方法。

Cloudant 的 DNA

Cloudant 出生并成长于云时代。它具有弹性和高可用性，可管理热门 Web 和移动数据类型，例如 JSON、全文本和地理空间数据。Cloudant 将最佳开源代码和领先见解融入为一些世界最大型 Web 和移动应用程序提供支持的创新 NoSQL 数据库平台。Cloudant 也定期向这些开源项目贡献代码。



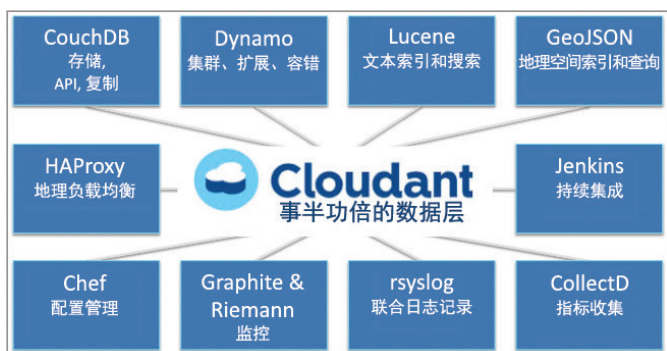


图 1: Cloudant 充分利用一流开源技术，提供与各类开发工具、平台和应用程序基础架构的兼容性

Cloudant 性能特点

Cloudant 最适合需要运营数据存储执行以下操作的应用程序：

- 处理大规模并发低延迟读写。
- 跨多个数据中心或设备复制和同步数据的可读、可写副本，为网络连接不可靠的远程或移动用户提供持续的数据可用性和离线数据访问。
- 管理多结构化或非结构化数据。
- 利用基于位置的服务（地理空间）、全文搜索和实时分析丰富应用程序。

以下表格提供了实际客户的真实性能数据特点：

大型单租户专用（单租户）客户

读写比率	20 : 1
数据量	130 TB
事务吞吐量	每天逾 20 亿项请求
延迟	低于 10 毫秒
用户群	全球（移动应用程序）
集群增长	12 个月内从 6 台服务器到超过 200 台， 横跨两个不同地理位置的数据中心

注：读写比率因客户而异。Cloudant 处理读取密集、写入密集或两者均衡的混合作业负载。

大型多租户客户

读写比率	3 : 1
数据量	215 GB
事务吞吐量	每天 100 万项请求（每小时 42,000 项）

这些是帮助审视 Cloudant Data Layer 性能的一般原则。随着应用程序需求的变化和工作负载的增长，Cloudant 的单租户（专用）集群可以通过添加节点来提升性能，从而与之保持同步增长。此选项按月收费，许多客户采用此选项实现“按需购买”。

Cloudant Data Layer 工作原理

Cloudant Data Layer 规模庞大；您的数据库在其内部而不是外部进行扩展，我们持续监控和扩展 Cloudant 以满足客户的整体存储和性能要求。

应用程序数据驻留在您创建的逻辑数据库中。Cloudant 将数据实际存储到服务器节点上，此节点加入到多租户或单租户集群中；Cloudant 将此集群作为在线服务进行管理，或者您使用 Cloudant Local 在自己的私有云中自行管理此集群。Cloudant 集群易于跨私有或公共云数据中心进行分配，让您可以创建全球“数据交付网络”并将数据推送到网络边缘，从而确保数据始终可用、访问迅速。

对于 Cloudant Managed 客户而言，Cloudant 自动处理负载均衡、集群、集群扩大/缩小以及系统高可用性等问题。随着您的数据库规模和用户负载的增长，Cloudant 会对数据层进行扩展，确保您跟上工作负载的步伐。Cloudant Local 客户负责所有上述管理任务，但却获得部署灵活性，可以将 Cloudant 安装到客户防火墙后面的客户基础架构上。

Cloudant 团队提供优化的开发工具和培训，帮助 Cloudant Local 客户确保集群持续运行和发展。

IBM Cloudant 入门

Cloudant 的入门和使用非常简单明了：

1. 创建 Cloudant managed DBaaS 帐户

Cloudant Managed DBaaS 让您省去了选择和配置硬件以及安装和配置 DBMS 软件和负载均衡器的繁琐过程。您只需在以下网址免费注册帐号：[Cloudant.com](https://cloudant.com)，您的数据层就准备就绪了。您将获得一个定制 URL，通过此 URL 可以立即开始创建数据库并处理您的数据。

激活 Cloudant 帐户后，您可以选择一个物理平台（例如 IBM® SoftLayer®）和地理区域（例如美国、欧洲、亚洲等）作为数据层的主要位置。通常情况下，选择取决于应用程序代码的存放位置，以便实现应用程序和数据层共置一处的优势。您可以随时将数据层从一个主机提供商或位置移动到别处，并有 Cloudant 数据专家 24x7 全天候监控您的数据，让您高枕无忧。

创建多租户与单租户帐户

默认情况下，您的数据层将驻留在一个同时被其他多租户帐户持有人使用的多租户集群中。多租户集群中的数据库安全可靠，其他 Cloudant 用户无法访问，除非您明确设置他人拥有共享权限以支持团队开发。Cloudant 包含特殊 IO 电梯，有助于确保您的应用程序性能不会受到“吵闹邻居”的影响。对于青睐专用数据层资源的用户，Cloudant 可以在来自多个数据中心甚至是云提供商的专用硬件上设置单租户集群，确保您的独特性能和可用性要求得到满足。

数据库即服务和 Service Level Agreement

Cloudant 的数据库即服务 (DBaaS) 解决方案让您可以购买可靠的数据管理“服务级别协议” (SLA)，而不是单纯购买数据库技术。这可以极大简化应用程序开发和交付。

2. 存储 JSON 数据

Cloudant 设计用于存储、索引和报告扁平的自述 JSON 文档集。如果您熟悉关系型 SQL 数据库，可以将 JSON 文档比作行，其中的字段比作列。Cloudant 中的“数据库”是 JSON 文档集。数据库内没有强制实施模式概念；因此，同一数据库内的 JSON 文档之间可能存在巨大差别。

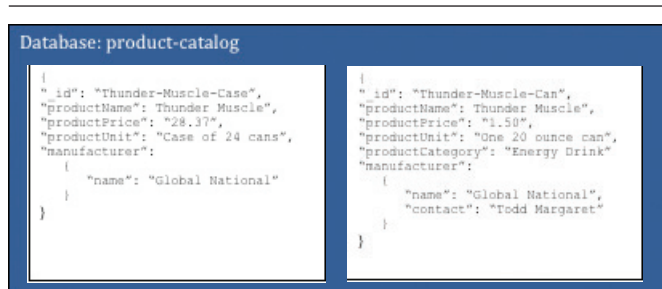


图 2：JSON 文档类似于关系数据库中的行，包含数值、字符串、日期、布尔、子结构、阵列和附件等字段类型

Cloudant 为每个 JSON 文档分配唯一的标识符 (`_id`)。文档字段可能是数值、字符串、布尔、对象、阵列、日期（JSON 字符串序列化/反序列化为 JavaScript 日期）或附件。元素数量或文档大小不存在限制。您可以创建索引并查询数据库中任何 JSON 文档的任何字段。

您可以根据需要向数据库安全地添加新文档类型，或修改现有文档的内容结构。与关系型数据库不同的是，这无需数据迁移或模式更新。这种灵活性非常适合支持敏捷的开发实践，消除数据库模式更新的复杂性，从而支持不断变化的代码库。这还让 Cloudant 非常适合处理结构多变的数据，例如产品目录、实验结果、内容管理信息，以及不是特别适合放入关系表的其他类型的记录。

使用 RESTful Web 服务 API，通过 HTTP 访问文档。本白皮书将在下文更详细地阐述数据库访问话题。

3. 建立数据索引

在关系数据库中，“索引”通常被视为性能优化。在 Cloudant 的 NoSQL 系统中，没有索引则将无法从数据库获取数据。

在 Cloudant 或 CouchDB 中，索引更类似于物化视图。事实上，索引是访问数据库数据的主要途径。所有数据库自动拥有一个基于各个文档预留 `_id` 字段的主索引。用户可以基于文档中的任何其他字段创建辅助索引，系统将根据 JSON 文档变更增量更新所有索引。和 CouchDB 一样，Cloudant 用户可以将辅助索引定义为使用 JavaScript 编写的 MapReduce“视图”函数。索引通常作为 RESTful API URL 路径的组成部分，查询据此运行。

值得注意的是，Cloudant 独特之处在于内置于服务中的额外索引系统，用于组织不同的数据类型并对传统 CouchDB MapReduce 视图系统的复杂性进行抽象：

Cloudant Query

Cloudant Query 是 Cloudant 的声明性查询系统，借用了 MongoDB 自己的声明性查询系统的语法。Cloudant Query 创建了一种数据访问语言，这对于多个 NoSQL 文档导向数据库而言都不陌生。

Cloudant Query 性能优于基于 JavaScript 的传统视图系统，因为 Erlang 中也部署了它。它的运行方式更加贴近 Cloudant 的核心数据库引擎，Erlang 也部署了该引擎。此系统还允许用户轻松定义一个索引并据此运行多个不同类型的查询，而不是定义多个视图来支持复杂查询。

可链接、增量 MapReduce

Cloudant 包含多种数据索引和查询方法。传统方法是在数据库中创建特殊“设计文档”，其中包含使用 JavaScript 编写的 MapReduce 函数。Map 函数定义哪个 JSON 文档归入索引，以及从每个成员 JSON 文档包含哪些字段。可选的 Reduce 函数定义依据索引中的数据执行的聚集操作，例如总和、平均值、计数等。每个设计文档都有唯一的 URL，从客户端应用程序访问此 URL 会将数据（或您使用 URL 中的可选查询参数定义的一些子集）作为 JSON 返回给客户端。

MapReduce 索引进行增量更新。Cloudant 把 MapReduce 函数分配到您数据库集群中的所有节点。对 JSON 文档做出的变更将引发增量索引更新，而不是重建整个索引。这可以降低 Cloudant Data Layer 的 IO 开销，还可以帮助您实施更快、更接近实时的数据分析。

为进一步改善性能并简化开发，Cloudant 帮助消除了关于只能有一个 map 阶段和一个 reduce 阶段的限制。“可链接”MapReduce 查询允许一个 MapReduce 作业的输出持续存留，并注入一连串的后续 MapReduce 作业中。这有助于更轻松、更快速地创建更多高级运营数据分析。

Lucene 索引和全文本搜索

为了处理全文本搜索，Cloudant 将 Apache Lucene 文本索引和搜索功能集成到 Cloudant 中。这可以帮助您借助以下功能丰富您的 JSON 数据索引和查询：

- 排序搜索 – 索引结果可以按照相关性或自定义排序字段进行排列
- 强大的查询类型 – 包括词组查询、通配符查询、邻近查询、模糊查询、范围查询等等
- 特定语言分析器
- 分面搜索和筛选
- 书签 – 以网络搜索引擎的风格对结果进行分页

基于 GeoJSON 的 2D 和 3D 地理空间索引和查询

GeoJSON 支持也集成到 Cloudant 中，用于支持开发基于位置的高精度服务，这种服务是移动和传感器网络应用程序的常见要求。多数其他 NoSQL 数据库仅支持邻近性和边界框等有限的地理空间处理功能，而 Cloudant 则允许您：

- 使用边界多边形、长方形、圆形或椭圆形创建查询
- 执行最近邻、路线优化和预测路径分析
- 存储广泛的数据类型，包括复杂的几何图形和元数据
- 使用坐标参考系 (CRS) 库，得益于 Cloudant 支持任何 CRS

4. 通过 RESTful API 访问 JSON 数据

通过 HTTP(s) 使用 GET、PUT、POST、DELETE 语句，可以单独或批量检索、存储或删除 JSON 文档。Cloudant 始终对传输中的数据加密。在 Cloudant 中，JSON 文档可以存储紧靠其他 JSON 数据类型的二进制附件，因而易于管理媒体文件，甚至代码。

示例：读取文档

Cloudant 的 RESTful API 让数据库中的每个文档都可以作为 JSON 通过 URL 进行访问；这是让 Cloudant 成为 Web 和移动应用程序的强大工具的功能之一。您可以使用命令行中的 curl 工具甚至是直接在浏览器地址栏访问这些端点。

此示例使用以下地址上的示例数据库（称作 `crud`，是 create、read、update 和 delete 的简称）：

`https://<accountname>.cloudant.com/crud/`。它包含单个文档，其中主文档 `_id` 是 `"welcome"`。

要通过 API 访问它，只需将此文档的 `_id` 添加到数据库的 URL 中，如下所示：

`http://<accountname>.cloudant.com/crud/welcome/`。在以下地址访问一个有效示例：
`http://broberg.cloudant.com/crud/welcome/`。

Cloudant API 兼容 Apache CouchDB。因此，您可以从适用于所有主要开发平台的 CouchDB 兼容客户端库、工具和框架访问 Cloudant。

5. 复制数据

多主复制功能支持 Cloudant 将数据的可读和可写副本分配到多个数据中心、设备和云提供商，并确保对副本所做的变更保持同步。这可以把用户连接到最近的数据副本，从而提高正常运行时间并降低访问延迟。它对易变应用程序数据进行的操作与内容交付网络 (CDN) 对静态内容进行的操作一样。

相反，其他热门数据库即服务产品和多数关系数据库采用“主-从”复制架构，其中只有一个副本可以进行更新（瓶颈），其他副本为只读。

您可以在您数据中心的 Cloudant Local 数据库和公共云平台上运行的 Cloudant Managed DBaaS 之间进行复制。您还可以在 Cloudant 和 Apache CouchDB（及其众多分支）之间进行复制，通过在分支机构、销售点和其他位置应用免费的开源数据存储，以更廉价的方式提高您的数据规模。

您可以有效控制复制内容、复制操作频率，以及当同一 JSON 文档被修改成多个副本时如何解决变更冲突。您可以定义过滤器，控制在何处复制什么文档的规则。您可以将复制设置为间断性执行（和频率），或设置为持续性执行以确保数据实时移动。Cloudant 所拥有的 API 让用户可以轻松查看数据库或帐号级别的变更、修改冲突，并制定冲突化解规则。

Cloudant 客户使用复制和同步执行以下操作：

- 出于容错目的而在多个位置创建数据的副本；如果一个源离线，还有多个源可用于响应请求
- 将数据推送到“边缘”数据库 – 例如数据集市或电子表格；使其非常适合独立的分析项目
- 通过中心辐射型数据库和移动同步支持离线计算
- 使用复制过滤器将数据子集分配到特定合作伙伴、分支机构等。例如，如果某个国家/地区的数据保护法要求来自或关于其公民的数据存储于该国家/地区境内，您可以对该国家/地区的 Cloudant 或 CouchDB 进行设置，以存储当地数据，并将其他数据的变更复制到该国家/地区境外的服务器。

6. 监控和可视化

为了让 Cloudant 及其客户监控数据层性能，我们定义了 API 和仪表盘来收集和报告系统指标。Cloudant 每秒生成 50,000 个指标。您无需成为理解此信息的专家，我们将为您代劳，确保您的数据库平稳运行和扩展。

如果您想亲自查看这些标准，可以通过 Cloudant 仪表盘和通过 API 访问您的帐户指标。

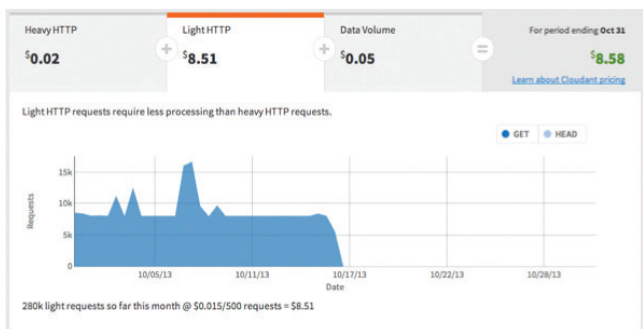


图 3: 通过 Cloudant 仪表盘可轻松访问数据标准

扩展 Cloudant 富有弹性的云可扩展性

为了满足当今应用程序的扩展要求，Cloudant 使用 Erlang（专门创建大规模并行应用程序的语言）实施了自己的无主、基于仲裁的集群版本，它可以应对：

- 集群成员
- 数据库交互的路由和协调
- 分布式查询的协调
- 远程过程调用，通过在远程节点上运行作业，实现更高的性能和节点故障恢复能力。

可调的最终一致性：

基于仲裁的集群可以帮助您指定您希望存储的数据副本数量（针对高可用性）、有多少副本必须写入磁盘，或必须有多少匹配才能让 Cloudant 认为数据已安全写入或具有安全的一致性。这些仲裁值被称作 N、W 和 R。

例如，在 3 节点集群中，您可以决定存储 N=3 个数据副本，每个节点一个副本。W=2 将意味着，在数据的两个副本写入后，Cloudant 将告诉您数据已安全提交。R=2 意味着，如果至少 2 个数据副本一致，则 Cloudant 将认为一个读取数据副本与其余一致。

仲裁值可以更改，以调整分区环境中的性能和一致性（例如 CAP 定理）。默认情况下，Cloudant 帮助优化可用性。

IO Queue (IOQ)

Cloudant 每天为数以千计的数据库处理数十亿项数据库交互。IOQ 是个成熟的 IO 优先级分层，用于分析各个数据库 IO 请求的优先级，以确保所有客户端获得公平份额的 IO 分层。它确保低延迟请求获得更高优先级，同时确保数据库压缩等较低优先级的 IO 获得较好的性能。

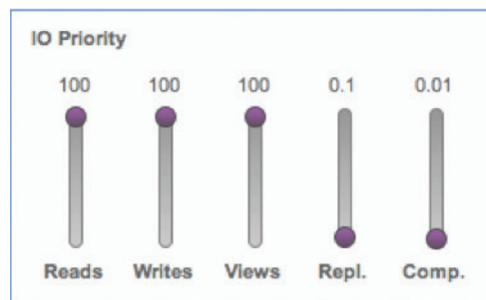


图 4: Cloudant 成熟的 IO 优先级层 (IO Queue) 会分析每个数据库 IO 请求的优先级

部署选项

Cloudant 提供多种形式：

Cloudant Managed DBaaS

- 一种可扩展、始终在线的分布式 NoSQL 数据层，由 IBM Cloudant 大型系统专家提供 24 x 7 全天候管理。
- 帐户可以运行在数据库服务器和负载均衡器的多租户数据库集群上，由许多用户共享，或
- 单租户集群，运行在由 IBM SoftLayer、Rackspace、Amazon AWS 或 Microsoft Azure 等顶级云提供商托管的专用硬件服务器上

Cloudant Local

- 将 Cloudant DBaaS 的功能融入您数据中心的私密环境中，并由您控制
- 客户借助基于六年管理型 DBaaS 运营经验而打造的 Cloudant 开发操作工具，利用自己的硬件在私有、公共或混合云平台上管理自己的 Cloudant 集群

私有云要求

Cloudant Local 融合 Cloudant 始终在线技术、简化的 DBaaS 操作工具，以及 Cloudant 开发和操作人员多年来积累的经验知识，将 DBaaS 的功能融入您数据中心私密环境中。

Cloudant local 要求概览

服务器节点

Cloudant Local 需要五 (5) 台机器打造功能齐全的 Cloudant 集群并确保 24 x 7 全天候可用性。

- 1 台主要负载均衡器
- 1 台故障转移负载均衡器
- 3 台数据库服务器

软件规格

Cloudant Local 运行于以下 64 位操作系统上:

Debian 衍生的 Linux 发行版:

- Debian 6.0.10
- Ubuntu Server 12.04.4

Red Hat 衍生的 Linux 发行版:

- Red Hat Enterprise Linux Server 6.5
- Community Enterprise Operating System (CentOS) 6.5

最佳实践是让所有服务器节点运行相同操作系统。

硬件规格

数据库 (DB) 节点

最低要求:

四 (4) 个内核和八 (8) 个线程, 例如 Xeon E3-1270 V2, 八 (8) GB RAM 和一个 (1) 千兆网络。

参考规格:

对于较大型部署, 最低要求是 12 个内核和 24 个线程 (例如双 Xeon E5 2620), 64 GB RAM, 用于满足应用数据卷要求的本地固态硬盘 以及一个 (1) 千兆网络。

负载均衡器 (LB) 节点

最低要求:

双核处理器和 4 GB RAM、500 GB 本地硬盘 和一个 (1) 千兆网络。

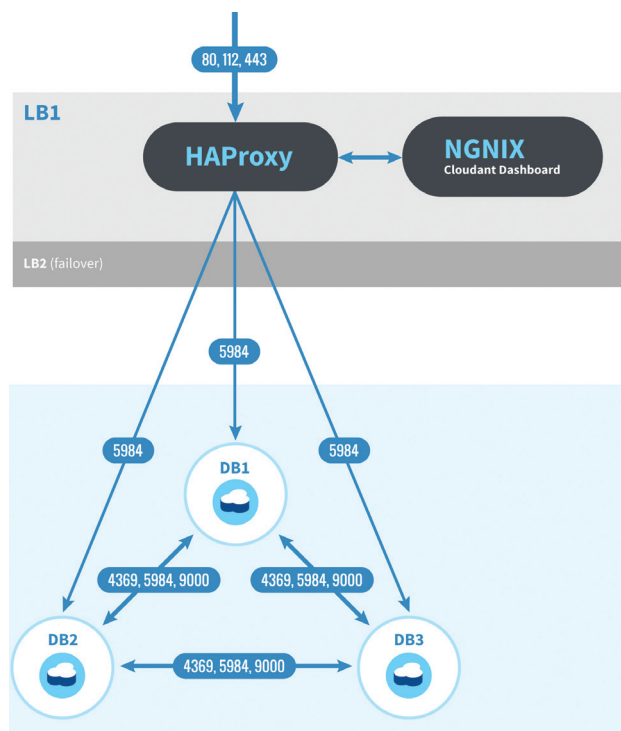
参考规格:

对于较大型部署, 最低要求是一个四核处理器和 8 GB RAM、1 TB 本地硬盘和一个 (1) 千兆网络。

Cloudant Local 架构示意图和端口

以下示意图显示的是 Cloudant Local 基本架构, 其中包括两个负载均衡器 (lb1 和 lb2) 和三个数据库节点 (db1、db2 和 db3)。

在此示意图中, 线段和附带文本说明了 Cloudant Local 组件所使用的通信路径和端口。



Cloudant 入门

Cloudant 提供全管理型 NoSQL 数据库即服务 (DBaaS) 解决方案, 用于快速、交钥匙式配置和无忧数据管理。它还提供 Cloudant Local, 将 Cloudant DBaaS 的功能融入您数据中心的私密环境中。您甚至还可以通过连接 Cloudant Local 与 Cloudant Managed DBaaS 数据库打造混合云数据库, 从而实现云成本、范围、性能和合规控制的平衡。只需在以下网址注册免费帐户即可开始入门内容的学习:

<https://cloudant.com>

更多详细信息

如需了解更多信息, 请联系您的 IBM 代表或 IBM 业务合作伙伴, 或者访问: cloudant.com/ 或 ibm.com/cloudant



© Copyright IBM Corporation 2015

IBM Corporation
Software Group
Route 100
Somers, NY 10589

2015 年 2 月

IBM、IBM 徽标、ibm.com 和 Cloudant 是 International Business Machines Corporation 在美国和/或其他国家/地区的商标或注册商标。如果这些和其他 IBM 带商标的术语在本文中第一次出现时带有商标符号 (® 或 ™)，则这些符号表示在本文发布时它们是 IBM 的美国注册或习惯法商标。此类商标在其他国家/地区也可能是注册商标或习惯法商标。IBM 商标的最新列表可在下述网页的“版权和商标信息”中查看：ibm.com/legal/copytrade.shtml。

Linux 是 Linus Torvalds 在美国和/或其他国家/地区的注册商标。

Microsoft、Windows、Windows NT 和 Windows 徽标是 Microsoft Corporation 在美国和/或其他国家/地区的商标。

其他产品、公司或服务名称是其他公司的商标或服务标记。

本文档为初始发布日时的最新文档，IBM 可能随时对其进行更改。并非所有产品/服务在 IBM 开展业务的每个国家/地区均有提供。

本文档中引用的性能数据和客户示例仅作参考用途。实际性能结果可能会有所不同，具体取决于特定的配置和操作条件。对于与 IBM 产品和程序配合使用的其他任何产品或程序，用户应负责相关的评估与验证工作。本文档中的信息“按原样”提供，不带任何明示或暗示的保证，包括不带任何适销性、对特定用途的适用性的保证，以及任何不侵权的保证或条件。IBM 产品根据提供这些产品时所依据协议的条款与条件进行保证。



请回收利用