



亮点

- 面向云环境的集成式硬件和软件平台性能卓越，易于部署，为用户提供强大支持。
- 借助内置的 IBM® Data Science Experience Local 创建高级分析模型。
- 加快分析工作任务的数据加载速度，缩短数据查询响应时间。
- 通过 IBM 的通用 SQL 引擎和内置数据虚拟化功能，在整个混合数据管理基础架构中实现无缝的互操作性。

IBM Integrated Analytics System

加速数据科学发展

企业是成为行业领导者，还是被竞争对手远远抛在身后，在这其中，能否运用分析能力加快决策过程是决定性因素之一。然而，为企业提供分析能力的过程会因为多种因素而变得复杂，其中包括：

- 务必让数据科学家参与整个过程，以便能够支持复杂的分析需求（包括机器学习）
- 必须支持企业现有的多种数据类型，包括非结构化和半结构化数据
- 物联网 (IoT) 等新应用的出现，要求实现实时分析和流式分析
- 呈指数级增长的数据持续制造数据“孤岛”，因此必须在企业内部环境和云环境中对这类“孤岛”进行有效管理
- 需要性能卓越、易于扩展的数据仓库基础架构，满足服务级别协议的要求

为了控制这些因素，同时保护现有投资，企业必须选择具备以下能力的分析和数据仓库解决方案：

- 为数据科学家提供所需的解决方案，最大程度消除开发和部署最先进的分析工作任务时所遇到的困难。
- 支持多个数据类型、位置以及传统数据仓库、近实时运营工作任务或混合工作任务。
- 适用于混合数据战略，并提供迁移到云环境的途径。
- 为任何规模或类型的工作任务提供高性能及可扩展性。
- 提供云环境所具有的灵活性，但部署在数据中心内。

投资的解决方案是否适当，对于企业未来能否扩展分析能力而言至关重要。选择是否适合，决定了企业是能够满足服务级别协议的要求，还是遭遇发展瓶颈。



IBM Analytics

解决方案简介

IBM Integrated Analytics System 提供经过优化、面向云环境的高性能数据平台，帮助数据科学家缩短开发和部署时间，深入挖掘所需的洞察，从而提升竞争力。

企业的数据科学家可以利用内置了 IBM Data Science Experience 的统一数据科学解决方案，在企业数据存储位置进行数据科学活动。这使得数据科学家能够开发更出色的机器学习分析能力，从性能经过优化的通用 SQL 引擎以及嵌入式 Apache Spark 处理中获益。

IBM Integrated Analytics System 经过精心设计，几乎不需要任何调优和维护，即可部署和管理各种工作任务，能满足要求最严苛的工作任务所需的超高性能和 PB 级可扩展性。IBM Integrated Analytics System 以系统中嵌入的 Apache Spark 处理引擎支持机器学习，从而实现更高性能的分析。该功能还有助于减少将分析和数据移至不同环境所带来的复杂性。IBM 混合数据管理产品系列中共享的通用 SQL 引擎能够使用企业现有的内部应用和云应用。这样的灵活性使您可以为特定任务选择合适的环境。

集成架构结合了软件增强功能，例如，非对称大规模并行处理 (AMPP)，还包含 IBM Power® 技术和闪存存储硬件。IBM Integrated Analytics System 可以处理传统数据仓库工作任务以及混合式运营工作任务。这些工作任务通常需要处理针对海量数据的查询、针对小数据集的快速精准查询，此外还需要处理多个并发操作访问。而 IBM Integrated Analytics System 可以通过单一解决方案，支持涉及多种数据类型和位置的广泛分析应用。这样的灵活性能为数据科学家带来无限可能。



借助 Spark 实现嵌入式分析和机器学习能力

IBM Integrated Analytics System 通过嵌入式 Apache Spark 处理能力，帮助数据科学家简化训练和评估预测模型的工作。通过这种功能，企业无需耗费大量时间将数据转移至其他系统以及转换为其他格式。数据科学家使用所选的工具开发模型后，就可以在数据存储位置进行测试、部署和训练等工作。由于每个节点都包含自己的 Spark 执行器进程，因此可以最大程度减少延迟，因此与单独的 Spark 集群相比，这有助于加快数据访问和计算速度。对于那些需要数据科学家从系统上卸下工作任务的情况，行业标准工具和通用 SQL 引擎能够将模型无缝移至 Spark 集群。

IBM Analytics

解决方案简介

这不仅能够简化流程，还可以为分析（包括机器学习）功能实现更卓越的性能和灵活性。数据科学家可以快速与系统中的数据对接，运用 IBM Data Science Experience Local 随附的五种经过授权的用户许可开始构建模型。这种基于云的互动式协作环境使数据科学家能够使用多种工具将洞察付诸实践。数据科学家还可以选择使用 Python、R 或 Scala，并可以通过系统中包含的 Jupyter Notebook 容器使用 Jupyter Notebook。Jupyter 可用于执行互动式代码，通过点击式部署，将代码转换为完成编译和部署的 Spark 应用。

除了提供用于数据挖掘、预测、转换、统计、地理空间数据和数据准备方面的预置功能外，IBM Integrated Analytics System 内置的 Spark 功能还支持开源 R 以及 Python、Java、C、C++ 和 Lua 等其他编程语言。

IBM Integrated Analytics System 还包含嵌入式 IBM Netezza® Analytics 技术以及多种算法，包括线性回归、决策树聚类、K 均值聚类和兼容 Esri 的地理空间扩展。该系统能够与业务分析及可视化工具协同使用，这些工具包括 IBM Cognos®、SAP BusinessObjects、Kognitio、Microsoft Excel、QlikView、SAS、Microsoft SQL Server Reporting Services (SSRS) 和 Tableau。该系统还可以处理模型构建和评分工具，比如 IBM SPSS®、Fuzzy Logix、开源 R 和 SAS。

集成与性能

IBM 凭借 20 多年来与成千上万来自各地区、各行业客户的丰富合作经验，打造卓越设计，帮助企业简化分析系统的部署和管理。对数据中心内的软件和硬件进行配置，使之能够协同工作，组成性能卓越的单一解决方案。只需花费数小时加载数据，无需创建数据库索引，也不必费力调优数据仓库或者在数据仓库运行后进行重新调优。

使用 IBM Integrated Analytics System 及其随附的 IBM Db2® Warehouse 技术的客户，能够立即发现整个 IBM 混合数据管理解决方案系列中使用的通用 SQL 引擎所带来的好处。IBM Db2 Warehouse 专为数据仓库和分析工作任务而设计。而通用 SQL 引擎采用动态的内存中列式技术，支持基于 IBM Db2 和 IBM BLU Acceleration® 技术的多种工作任务。BLU Acceleration 大规模并行处理 (MPP) 架构旨在用于快速深入地分析海量数据，规模可达 PB 级。与先前系统相比，查询响应速度快了 100 倍，¹ BLU Acceleration 列列表可以和传统行式表共存于相同的模式、存储器和内存中，因此可以同时查询行式表和 BLU Acceleration 列列表。BLU Acceleration 技术使传统的内存中功能如虎添翼，使性能更上一层楼，即使数据集超过内存大小，也不会有任何问题。BLU Acceleration 的动态内存中列式技术以及数据忽略功能提供了一种扫描和查找相关数据的有效方法，即使在数据压缩的情况下也不受影响。

IBM Integrated Analytics System 利用 IBM Power Systems™ 和 IBM FlashSystem® 技术，提升硬件的可靠性和性能。目前的 IBM Power 架构支持更密集的系统，只需更少的节点就能获得和先前产品相近的性能。作为系统的默认存储，IBM FlashSystem 可实现超低延迟、接近内存水平的高速 I/O 以及出色的可靠性。

分析应用以峰值性能持续运行，此外，IBM Integrated Analytics System 还提供更高水平的可靠性，帮助您满足或超越服务级别协议的要求。由于采用容错设计消除单点故障，因此 Power Systems 和 IBM FlashSystem 存储产品的额定正常运行时间显著延长。2017 年的一项信息技术情报咨询 (ITIC) 调研结果显示，IBM Power Systems 在任何主流 Linux 服务器平台上的计划外宕机时间都是最短的（每台服务器每年 2.5 分钟）。²

IBM Analytics

解决方案简介

整个系统中的所有组件都实现冗余，以便确保在硬件发生故障时运行不会中断。该系统还包括额外的内置高可用性，用于提供自动故障转移功能，实现持续的高性能表现。由广泛应用于 Db2 产品家族的 IBM Data Server Manager 支持的内置控制台监控和管理所有组件(包括硬件和软件)。

面向云环境和数据的混合方法

“一体适用”式的方法通常无法满足企业的数据处理要求。IBM Integrated Analytics System 基于 IBM 混合数据管理产品家族所使用的通用 SQL 引擎（一系列共享组件和功能），能够在整个基础架构中实现无缝的互操作性。

例如，团队所使用的数据仓库可能需要迁移至云端，以满足季节性容量需求。如果采用 IBM Bluemix® Lift 等工具，就可以将此工作任务无缝迁移至 IBM Db2 Warehouse on Cloud。通用 SQL 引擎可以确保您无需重写任何应用。

通用 SQL 引擎始终能够提供数据视图，无论数据位于何处，也无论是非结构化数据还是半结构化数据。通用 SQL 引擎中的系统内置数据虚拟化服务有助于统一整个逻辑数据仓库中的数据访问，支持联合 Db2、Hadoop 甚至第三方数据源。

可扩展性与扩展选项

软件和硬件架构都能够随着工作任务的增加而不断扩展，以支持业务发展的需要。计算容量和存储容量可单独扩展，具备云平台那样的灵活性和可扩展性。硬件扩展时无需中断业务，可直接在系统运行时进行。

IBM Integrated Analytics System 还支持根据数据的使用频率进行分层存储，确保最高水平的性能，即使数据量非常庞大，也不在话下。系统直接在存储节点上管理最近使用和非常活跃的“热”数据，而在更具成本效益的高密度 IBM Storwize® 存储设备上管理存在时间较长、不太活跃的“冷”数据。

应用

IBM Integrated Analytics System 采用的通用 SQL 引擎可确保特定工作任务在合适的部署平台上运行，同时确保数据始终可访问，无论其类型、位置或大小。以下是一些应用，您或许可以从中获得一些启示：

- 通过使用所有参与交互的 Hadoop、数据集市或其他相关部署创建逻辑数据仓库，提供内部和云端数据的统一视图，令卓越运营成为新常态。
- 借助内部数据以及分析功能，实时营造个性化的客户体验，提供用户所需的高性能和简便的可扩展性。
- 通过使用嵌入式机器学习加速开展复杂的分析，揭示有关用户的深入洞察，从而加快创新产品的上市速度。
- 借助能够快速加载数据而且只需很少或根本无需调优、也不需要大量配置工作的系统，更快地为业务用户提供所需的工作任务。
- 逐步增加混合式运营工作任务，扩展可供业务用户使用的工作任务选项。
- 以更准确的数据视图，帮助确保运营合规。
- 使用数据虚拟化技术，构建针对各种不同数据类型和数据集的分析应用，统一涵盖整个逻辑数据仓库和其他技术的数据访问，从而获得更高质量的洞察，促进业务发展。

产品规格

IBM Integrated Analytics System 集成并优化了所有计算、存储和网络资源以及分析和数据仓库软件。它以机架配置形式提供，如下表所示。

规格 - 单机架系统			
IBM Integrated Analytics System 型号	M4001-003	M4001-006	M4001-010
活动节点	3	5	7
CPU 核心数	72	120	168
内存 (TB)	1.5	2.5	3.5
用户数据 (TB) (假设 4 倍压缩)	64	128	192
每机架最大功率 (瓦)	4,400	6,900	9,400
散热能力 (BTU/小时)	15,000	23,500	32,000
机架重量 (千克)	479.9	563.8	647.7
高度 (厘米)	202		
长度 (厘米)	110		
宽度 (厘米)	64.8		
电源线	北美和日本 - 200–240 伏, 单相, 30 安 欧洲、中东和非洲 - 220–240 伏, 单相, 32 安		
托盘数/机架	4		
安全标准	UL/CSA 60950-1, 以及 EN 60950-1 和 IEC 60950-1		
排放标准	FCC part 15, ICES-003, AUS/NZ CISPR 22, VCCI, EN55022/55032 Class A ; European immunity : EN 55024、EN 61000-3-2 和 EN 61000-3-3		
注	可采用多机架系统形式提供。要了解更多信息, 请联系您的销售代表或 IBM 业务合作伙伴。		

关于 IBM 混合数据管理解决方案

IBM 提供了最全面的混合数据管理解决方案系列之一, 旨在帮助企业最大程度发挥信息资产的价值, 发掘新锐洞察, 以支持更快做出更明智的决策, 优化业务成果。

了解更多信息

要了解有关 IBM Integrated Analytics System 的更多信息, 请联系您的 IBM 代表或 IBM 业务合作伙伴, 或者访问我们的网站: ibm.biz/BdjfvP。

此外, IBM 全球融资部可以帮助您以最经济和最具战略意义的方式快速获得业务所需的软件功能。IBM 能够为信誉良好的客户量身定制符合其业务和发展目标的支付解决方案, 帮助他们实现高效的现金流管理, 平衡成本和效益, 获得更丰厚的投资回报。IBM 全球融资部能够为您所有关键的 IT 业务提供资金支持, 祝您推动业务发展。要了解更多信息, 请访问: ibm.com/financing。



© Copyright IBM Corporation 2017

IBM Corporation
IBM Analytics
Route 100
Somers, NY 10589

美国出品
2017 年 9 月

IBM、IBM 徽标、ibm.com、BLU Acceleration、Bluemix、Cognos、Db2、IBM FlashSystem、Power、Power Systems、SPSS 和 Storwize 是 International Business Machines Corp. 在全球许多司法管辖区的注册商标。其他产品和服务名称可能是 IBM 或其他公司的商标。Web 站点 www.ibm.com/legal/copytrade.shtml 上“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。

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

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产品或服务。

本文讨论的性能数据均在特定操作环境中测得，实际结果可能不同。用户负责评估和验证与 IBM 产品和程序一起运行的其他任何产品或程序。

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的协议条款和条件获得保证。

客户负责确保遵守任何适用的法律和法规。IBM 既不提供法律建议，也不表示或保证其服务或产品能确保客户符合任何法律或法规。

良好安全实践声明：IT 系统安全是指通过预防、检测和响应企业内部和外部的不当访问来保护系统和信息。不当访问可能导致信息被篡改、毁坏或挪用，或者可能导致您的系统被损坏或滥用，包括用来对他人进行攻击。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。

1 <http://www.redbooks.ibm.com/technotes/tips1204.pdf>

2 <http://itic-corp.com/blog/2017/06/ibm-lenovo-servers-deliver-top-reliability-cisco-ucs-hpe-integrity-gain/>

