



要点

- 利用多租户框架简化 Apache Spark 部署和管理
 - 通过高效资源调度和共享的基础架构，缩短实现成果的时间
 - 借助动态的分配，缩减成本并提高资源利用率
 - 集成 Spark 与其他应用程序，例如 Hadoop、Cassandra 和 MongoDB
 - 维护并简化多数据源的管理
 - 消除多个实例和不同版本的 Spark 带来的资源孤岛
 - 利用基于角色的访问、Kerberos 身份验证和 SASL 及 SSL 加密增强安全性
 - 自动化、动态的云爆发
-

IBM Spectrum Conductor with Spark

简化 Apache Spark 部署，缩短实现成果的时间并最大程度提高资源利用率

Apache Spark 是一种开源的大数据分析框架，具备无与伦比的性能优势。不过，Spark 的实施面临严峻挑战，包括投资新的专业技术、工具和工作流以及与其他框架集成。设置临时的 Spark 集群可能导致资源利用率低下，同时还会产生管理和安全问题。

IBM® Spectrum Conductor™ with Spark 可让这些问题迎刃而解，帮助用户克服部署和管理 Spark 的挑战。与其他需要零散组件装配的竞争性开源产品的不同之处在于，IBM Spectrum Conductor with Spark 是集成式解决方案，同时还有 IBM 服务与支持作它的坚强后盾。它既包含 Spark 发行版，又支持 Spark 和其他框架的多租户，同时还有细化动态资源分配技术的加持。这些技术已在许多要求严苛的客户环境中被成功广泛实施，可显著提升性能与效率。

组织可利用 IBM Spectrum Conductor with Spark 积极高效地部署 Spark 应用程序。这种企业级的多租户管理解决方案可支持多个 Spark 实例，能够实现资源利用最大化，增强性能与规模，并消除单独实施 Spark 所产生的资源孤岛。IBM Spectrum Conductor with Spark 支持集成 Spark 与其他应用程序框架，例如 Hadoop、Cassandra 和 MongoDB。



加快实现成果的速度

IBM Spectrum Conductor with Spark 能够在单个共享的基础架构上支持多个同时运行的 Spark 和其他框架实例，让应用程序充分利用各种可用资源。一种有口皆碑的高效资源调度器，能够进行精密的资源分配，实现非凡的应用程序性能，提高资源利用率并更快响应业务关键型需求。在运行多个应用程序工作负载的环境中，IBM Spectrum Conductor with Spark 分配的资源既能符合服务水平，同时还可在应用程序实例之间形成安全隔离。

IBM Spectrum Conductor with Spark 提供比竞争性开源资源管理器高多达 58% 的 Spark 作业吞吐量。¹ 它还提供先进图形处理单元 (GPU) 支持，可充分利用自动化调度和监控的全部力量，以改善利用率和管理。此外，缓存或永久的资源定义数据集 (RDDs) 可以跨用户共享，以避免重新加载或重新计算此前的结果。所有这些因素结合在一起实现了最快的成果实现速度，同时最大限度降低计算基础架构开支。

提高资源利用率

IBM Spectrum Conductor with Spark 有助于组织杜绝集群无序扩张，避免出现资源利用率低下的情况。通过在单个共享平台上运行工作负载，该解决方案能够让单个应用程序使用通常专用于其他应用程序实例并可能会被闲置的资源。

IBM Spectrum Conductor with Spark 还可支持多租户，用户可以在共享环境下同时运行多个实例和不同版本的 Spark。这种功能可以让各小组运行不同版本的 Spark，但无需步调一致地对这些版本进行升级，从而有助于组织管理 Spark 快节奏的生命周期。

当临时需要额外的资源时，Spectrum Conductor with Spark 可以利用它的资源连接器自动、动态地利用云资源。云爆发动态地发展和缩小您的集群，同时支持组织内部主机和云主机的混合搭配。当您的工作负载超过配置的阈值，在云资源上创建额外的主机。当云资源上运行的工作负载低于配置的阈值，主机将被关闭，从而管理并最大限度减少对云资源的利用。

降低管理成本

IBM Spectrum Conductor with Spark 提供先进的服务编排和工作负载管理功能，有助于抑制基础架构和管理成本。一款基于策略的成熟资源管理器，可实现动态资源分配，让组织能够优化现有硬件利用，减缓对增量资本投资的需求。管理员可利用统一的界面管理多个 Spark 框架，无需从各个框架逐一收集并整合各项指标。

轻松实施完整的解决方案

组织希望转向可优化信息资产存储、分析和保护的解决方案。IBM Spectrum Conductor with Spark 是一款集成解决方案，包含面向数据分析、工作负载管理、监控、报告和企业级安全性的 Spark 发行版。在存储管理方面，IBM Spectrum Conductor with Spark 可与 IBM Spectrum Scale™ 结合使用，后者的存储效率遥遥领先于 Hadoop 分布式文件系统 (HDFS)。若用户更愿意选择 HDFS，IBM Spectrum Conductor with Spark 也可提供支持。无论是对于探索性项目还是在生产环境中，内含的 Spark 发行版让框架易于部署。

轻松部署 Spark

IBM Spectrum Conductor with Spark 提供卓有成效的解决方案，可满足组织的种种诉求：

- 在 Spark 实例间共享资源，缩短实现成果的时间并提高利用率
 - 管理短暂的 Spark 生命周期
 - 为多个 Spark 实例提供精细的资源分配控制
 - 整合多个 Spark 环境并将 Spark 与其他框架集成
 - 针对计算密集型任务充分利用 GPU 处理力量
 - 在混合云环境中部署 Spark
-

为何选择 IBM?

IBM Spectrum Computing 推出了全面的软件定义基础架构解决方案产品组合，有助于组织以最高效的方式交付 IT 服务，优化资源利用率以缩短实现成果的时间并降低成本。这些产品有助于最大程度发挥基础架构的潜能，无论规模大小，都能加快分析、高性能运算（HPC）、Apache Hadoop、Spark 以及原生云应用程序，从数据中获取洞察，以更快的速度将更优质的产品推向市场。

无论是部署在数据中心还是云端，IBM Spectrum Computing 解决方案都能以更快的速度完成产品开发，制定关键业务决策并获得突破性洞察，该方案适用于金融服务、制造、数字媒体、油气、生命科学、政府、研究与教育等领域。从设计一级方程式赛车到信用风险分析，各行各业的组织都在使用 IBM Spectrum Computing，将它作为软件定义基础架构解决方案的基础，应对在大数据、分析、HPC 和云方面的挑战，从而获得更优异的业务成果。

如需更多信息

如需了解有关 IBM Spectrum Conductor with Spark 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：

- ibm.com/us-en/marketplace/spark-workload-management
- ibm.com/storage/software-defined-storage

此外，IBM 全球融资部提供多种付款选项，可以帮助您获得发展业务所需的技术。从采购到处理，我们提供 IT 产品和服务全生命周期管理。有关更多信息，请访问：

ibm.com/financing



© Copyright IBM Corporation 2017

IBM Systems
New Orchard Rd
Armonk, NY 10504

2017 年 11 月

IBM、IBM 徽标、ibm.com、IBM Spectrum 和 IBM Spectrum Scale 是 International Business Machines Corp. 在全球许多司法辖区的注册商标。其他产品或服务名称可能是 IBM 或其他公司的商标。有关 IBM 商标的最新列表，请访问以下网站的“版权与商标信息”部分：

ibm.com/legal/copytrade.shtml

本文档为初始发布时的最新文档，IBM 可能随时对其进行更改。

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 根据提供产品时的协议条款与条件提供产品担保。

¹ 《STAC 报告：IBM Platform Conductor for Spark、Apache YARN 和 Apache Mesos 的 Spark 资源管理器比较 — 第 1 阶段》，Securities Technology Analysis Center，2016 年 3 月 28 日，
<https://stacresearch.com/news/2016/03/29/IBM160229>



请回收再利用