

使用 MapR 发行版和高速以太网提供功能强大的分析



背景

移动技术和物联网 (IoT) 传感器在汽车、工业设备和医疗设备等领域的兴起，导致产生了无限丰富的数据。这些数据日趋呈现半结构化和非结构化的性质，并且通常存储在 Hadoop 和 NoSQL 系统中。对这些数据分析得出新的有用结论为成功的公司带来了业务增长和竞争优势。然而，为了让大数据发挥其巨大潜力，技术必须就位，使组织不仅可以捕获和存储数据，还可以快速获得新的业务见解，通过多种方式帮助提高业务绩效。利用 MapR 和 Mellanox 的联合解决方案，数据科学家可以更快地分析大量多样化数据集，并快速而有效地制定新的业务假设。

联合解决方案要点 MapR 融合数据平台

作为 Hadoop 领域的技术领导者，MapR Technologies 提供了一个融合平台，其中包括 Apache Hadoop 的完整发行版。在主要的 Hadoop 发行版

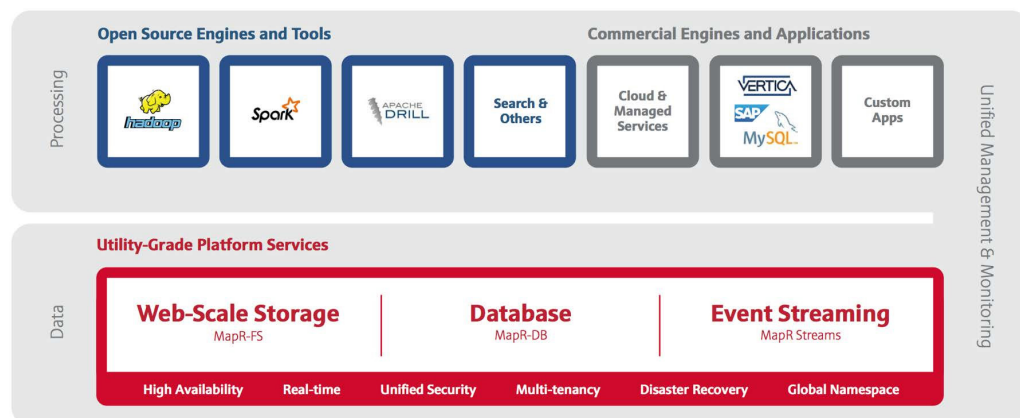
主要业务优势

- 执行分析和制定新假设的速度比以前更快
- 使用零接触部署持续运营
- 凭借 MapR 和 Mellanox 的企业级可靠性加速 Hadoop 部署
- 通过最大限度地利用系统资源和支持多个工作负载，以最经济的方式构建 Hadoop 集群，实现最低总体拥有成本

中，MapR 继续保持最快的速度，并且在特性和功能方面保持领先。MapR 融合数据平台将 Hadoop 和 Spark 与全球事件流、实时数据库功能和企业存储集成，用于开发和运行分析应用程序。MapR 平台服务层包括 MapR 文件系统和 NoSQL 数据库（分别为 MapR-FS 和 MapR-DB），提供了多种企业级功能，包括高可用性、数据安全性以及与 NFS 的轻松集成。该平台支持处理存储在文件以及 NoSQL 数据库中的结构化和非结构化数据的广泛应用程序。这种用于所有大数据工作负载的单一数据平台为硬件使用提供了最佳投资回报率。

"MapR 融合数据平台集成了 Hadoop、Spark、NoSQL 和流的强大能力，用于开发和运行创新的数据应用程序。MapR 和 Mellanox 之间的这种合作关系帮助客户从其大数据部署中获得更多的产出和价值。"

- Steve Woledge ,
MapR Technologies
产品营销副总裁



优化性能的高度可扩展平台

相对之前使用的大数据应用，MapR 的高性能数据平台使业务分析人员能够更快速地得出新的业务价值。这些工作负载可以是基于 SQL（使用 Hive、Impala、Spark SQL 等）的应用程序、基于 NoSQL（使用 MapR-DB）的新型应用程序，或使用机器学习、图形和流分析的更高级应用程序。由于 MapR-FS 的优势以及生产级 Hadoop 环境中闪存使用量的增加，网络带宽往往成为数据引入和分析过程中的瓶颈。如果性能受到网络的限制，那么为 MapR 投资购买更快的服务器和闪存存储就没有意义。Mellanox 为大数据应用程序提供最高效的端到端 10/25/40/50/56/100GbE 以太网网络。Mellanox 以太网交换机提供一致的低延迟，并且可以支持各种无阻塞、无损的架构设计。此外，使用 ConnectX®-3 Pro 网络适配器中基于硬件的无状态卸载和流转向引擎，利用基于 IP 的 TCP 或 UDP 传输的大数据应用程序能够实现最高的吞吐量和应用程序密度。这些高级无状态卸载可以降低 IP 数据包处理中的 CPU 开销，从而在大数据集群中以较少的时间完成更繁重的分析工作负载。

简化端到端管理

在大型横向扩展软件生态系统（如 Hadoop 和 NoSQL）中，在 MapReduce 应用程序以外部署更广泛的大数据软件生态系统（如 Spark、Drill 和 Sqoop）时，实现集群的高可用性和低停机时间成为数据中心管理员面临的关键挑战。MapR 凭借其他 Hadoop 发行版所不具备的简化业务连续性、分层存储、安全性和预配置功能，增加了管理的便利性。MapR 控制系统是基于浏览器的管理工具，管理

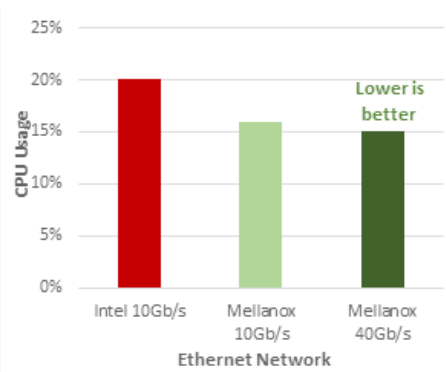
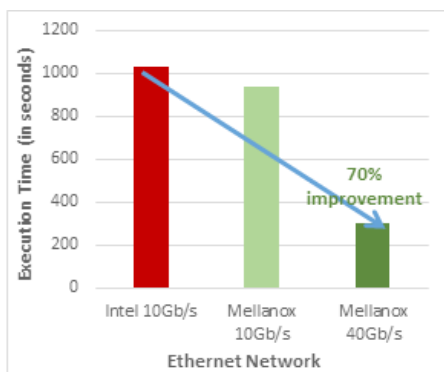
联合解决方案优势

- 使用更快的网络和闪存存储提升现有基础架构的性能
- 确保最高级别的可靠性以运行任务关键型环境
- 通过简单的配置简化数据体系架构

员利用该系统可以管理和监控 MapR 部署，以确保实现可靠、高性能的系统。Mellanox NEO™ 为 IT 管理员提供最先进的网络编排、自动化和监控平台，可获得整个网络的全景视图。此外，通过 NEO 中的一键功能，数据中心管理员可以在简单模板中轻松配置大型横向扩展网络，从而允许在集群中添加新数据节点时自动复制网络状态。

结论

虽然 Apache Hadoop 为分析大型和多样化数据集提供了强大的工具，但部署成功、可靠和高性能的基础架构可能是艰巨的挑战。MapR 的 Hadoop 发行版提供了关键的技术进步，使 Hadoop 实施可以轻松、可靠、快速地完成生产部署。通过利用 Mellanox 先进的大数据卸载功能，IT 组织不再需要在性能和可靠性之间作出取舍。凭借这些集成功能，企业可以更快地实现大数据分析的竞争优势，并且风险更低，还有信心实现企业级生态系统。



北京市朝阳区望京东园七区保利国际广场 T1 15 层
Tel: 010-5789 2000
www.mellanox.com