

有助于实现高性能数据与 AI 架构的高效提示指南

通过基于高性能基因测序与成像真实用例的提示，了解如何准备数据，实现精准医疗、体验新水平的速度和规模。

“大海捞针”

每个医疗保健和生命科学组织都希望转型为数据驱动型组织，实现突破性的科研，以支持新的诊断与治疗方法的开发。

但是如何做到这一点？

若要成为数据驱动型组织，组织需要实施全局式方法，以帮助他们：

- 解决数据数量、多样性和复杂性不断增长的问题，这些数据都分布在各个独立的系统中而且需要进行集成。
- 实现整个工作流的自动化，以分析数据并从中发掘价值，同时支持快速演变的框架和应用。

从 IT 的角度来说，这就需要新水平的速度、可扩展性、灵活性和高级管理功能，以支持协作、降低成本并加速获取洞察力。

本指南列出了 IT 架构充分利用数据价值、将海量数据集转化为可执行医疗洞察力所需的关键功能。部署强大的高性能数据与 AI (HPDA) 架构，有助于您的组织实现生物医疗研究转型并推动精准医疗的发展。



高效提示 1:

构建数据中心来管理海量数据

HPDA 数据中心应具备的五大关键功能:

数据生成

每个仪器都会生成数据而且仪器的速度会越来越快、规模越来越大，进而生成了需要进行存储并快速处理的海量数据，其中大部分为非结构化数据。部署具有一流可扩展性和性能的解决方案，存储并分析由各种来源快速生成的海量数据，避免数据准备成为瓶颈。



数据分层

基于具体存储需求（需要高性能存储还是需要长期保留归档），以最快的速度将数据移至最优存储层，实现数据放置的自动化。这有助于针对现代化工作负载构建一个通用的存储池，确保对文件存储和对象存储的访问，进而降低资本开支和运营开支。



数据共享

为各个地域上分散的组织提供单个全局名称空间，让研究人员能够以最低的时间延迟、安全的方式共享、访问和分析数据，进而实现协作与生产效率的最大化。如此一来，便可确保随时随地将数据提供给需要数据的人员。



数据对等处理

复制当前的数据，使其能够在任何位置（远程、云端或其他站点）进行持续访问，无需制作大型数据副本。这种战略可帮助您避免数据重复带来的低效和高成本问题。



数据分类

跟踪快速增长且分布在许多独立位置和项目中的不同类型的数据以及这些数据的整个生命周期管理。对海量数据集进行标记和分类，使您可以快速、轻松地找出所需的准确数据。借助一个简单的仪表板，您便可高效地使用数据进行分析。



高效提示 2:

构建统筹安排程序，确保应用的妥善保护

HPDA 统筹安排程序应具备的五大关键功能:

并行计算

部署智能化工作负载和策略驱动型资源管理，以优化整个数据中心、内部环境和云环境中的计算集群，进而提升获取洞察力的速度并改善成效。如此一来，数百个并发用户便可在数千个节点上并行运行数百万个任务，既不会造成宕机，也不会导致用户或应用出现中断。



平台即服务

通过适于医疗保健和生命科学领域关键工作负载（基因测序、成像、临床诊断）的分析框架提升高性能计算基础架构，比如以平台的形式来使用 SPARK，以分析数据并生成数据集成所需的资源。



云计算

以弹性的方式将内部工作负载移至云端，以便您解决不可预测的需求并避免出现可能会导致一些任务运行变慢的瓶颈。通过多云工作负载统筹安排，您可以更快地获得洞察力并节省成本。



管道构建

基于各种不同的应用构建一个基因测序或成像管道，并将这些应用集成到复杂的工作流中，您可以通过共享式计算基础架构高效管理这些工作流，进而缩短实现成效所需的时间。



应用中心

通过灵活且易于使用的图形用户界面工具，让集群用户和管理员能够高效管理环境。通过该工具，您可以实时了解应用状态，并通过直观的、可自归档的标准化界面与应用进行交互。



高效提示 3:

实现数据中心与统筹安排程序之间的协同：数据中心 + 统筹安排 = 高性能、低成本、易用性、协作

按工作负载进行多云统筹安排

将数据放置在最能在组织运行其应用和服务时发挥数据价值的位置，让计算资源能够“感知”数据。当某个计算工作负载被分派到另一个位置（如不同的站点或公有云）时，统筹安排程序便会与数据中心协同工作，实现数据的同步移动。



零宕机

让计算资源能够“感知”IO，进而防止 IT 基础架构出现过载，甚至是崩溃。在基础架构过载之前，数据中心便可检测到这一趋势并与统筹安排程序协同工作，平衡并控制工作负载的分派，以实现最大化并确保业务连续性。



采取下一步行动 - 了解可用资源

观看该网络研讨会，了解世界级组织如何借助这些提示推动实现精准医疗：[网络研讨会：体验新水平的速度和规模：高性能基因测序和成像](#)

了解我们基于软件定义基础架构为医疗保健和生命科学行业设计的高性能数据与 AI 参考架构：[视频：面向医疗保健行业的高性能数据架构](#)

免费试用！

[试用 IBM® Spectrum Scale™](#)：立即开始 30 天免费试用，体验一流的可扩展性以及如何降低存储成本

[试用 IBM Spectrum™ LSF](#)：确保将适当的资源自动分配到适当的任务，以降低成本并实现性能最大化：



有关更多信息及相关案例研究，敬请访问：

面向医疗保健行业的数据存储

ibm.com/it-infrastructure/storage/healthcare-industry

© Copyright IBM Corporation 2018
IBM Corporation
IBM Systems
Route 100
Somers, NY 10589
美国印刷
2018 年 9 月

IBM、IBM 徽标、ibm.com、Spectrum 及 Spectrum Scale 是 International Business Machines Corporation 在世界各地司法辖区的注册商标。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 上的“Copyright and trademark information”部分中包含了 IBM 商标的最新列表。本文档截至最初公布日期为最新版本，IBM 可随时对其进行修改。IBM 业务合作伙伴设定了自己的价格，而且这些价格可能会发生变动。IBM 并不一定在开展业务的所有国家或地区提供所有这些产品或服务。

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的条款和条件获得保证。

81019181-CNZH-00