



生命科学高性能计算 解决方案

生命科学是计算机、生物、数学等多学科交叉而新兴的学科分支，以借助计算机科学、信息科学等领域的算法与工具解决生命科学中的问题为主要特征。

随着新一代测序技术的发展，生命科学迎来大数据爆发的时代。动植物生理、群体遗传学、进化生物学、分子动力学、结构生物学、流行病学等等，生命科学的各个专业都离不开测序数据的产生、处理。巨大的数据量处理能力和并行计算需求伴随当今生命科学的进展，因此高性能计算机在生命科学研究当中应用广泛且扮演着重要的角色。

■ 生命科学领域中HPC的重要性

- ◆ 测序技术飞速发展，基因序列数目呈指数级增长，对于数量庞大的基因进行同源性搜寻、比对、分析和遗传发育分析等，需要高性能计算巨大的数据处理量和并行计算量
- ◆ 生命科学研究对象往往是蛋白质和DNA大分子，对这些分子三维结构的预测，动力学特性、热力学特性在生命过程中如何发生作用，也需要借助高性能计算
- ◆ 高性能计算还可用于计算生物学、系统生物学、结构生物学、临床组学数据的分析存储、技术开发，辅助公卫中心进行病毒筛查、毒株分离、疫苗研发等工作

■ 生命科学领域建设高性能计算集群的难点

海量分子生物信息数据的出现，为生命科学研究提供了空前丰富的数据资源，同时也为计算带来新的挑战：

1. 数据计算方面

序列对比、基因识别、结构预测、序列拼接、分子进化、比较基因组学和药物设计等生物信息常规分析具有大规模、高通量、密集型特征，经常需要处理千兆乃至万兆数据，需要海量计算资源，且本地计算成本太高，扩容慢，收效更慢。

2. 数据存储方面

由于基因测序技术的提升，测序成本降低，从最初的人类基因组开始细分扩展，目前已经涉及到肿瘤、遗传病检测，扩展到植物、远古生物、细菌、病毒、微生物的基因检测。因此数据种类和数据量异常庞大，经常以PB为单位保存。

3. 数据安全方面

生物信息数据相对比较隐私，但传统的生命科学公司IT能力较弱，安全措施不到位，防御能力较弱。

■ 联智科技生命科学领域HPC解决方案概述

面向生命科学领域的HPC解决方案充分考虑生命科学领域应用软件特点，基于联智科技自主研发的CHESS高性能计算管理平台和二十多年行业经验的积累，为用户设计、定制、搭建和运维生命科学领域HPC集群系统，提供高效率、高带宽、低延迟的I/O优化服务，统一管理生物信息数据，提高计算速率，实现高并行访问和高数据共享，保证HPC集群高效稳定运行、按需调度、有效管理，构建一整套高可靠、高性能、高易用的生命科学计算平台，满足当前绝大多数生命科学领域对高CPU计算、大容量内存、高IO吞吐、高数据交互的计算需求，帮助生命科学领域机构高效率、低成本地实现科研计算需求。

■ 联智科技生命科学领域HPC解决方案价值

1.真正以生物信息领域应用为核心的集群建设服务

提供从系统层到应用层的一站式解决方案和全流程技术服务，保证用户集群高效部署、即买即用。

2.充分考虑生命科学领域应用软件特点

管理与规划海量测序数据，合理设计集群、优化软硬件配置，保证应用性能达到最优。

3.更高效地利用超算强大计算力

确保海量影像数据的存取与分析、异构加速处理与机器学习预测，实现系统全流程拟合与预测。

4. 科研成果的视图可视化

借助易用的管理平台和直观的图形界面实现多种应用软件的管理和使用，提高集群运行效率，简化集群运维管理。

5. 安全可靠、稳定易用的高性能计算平台

整套系统的高CPU计算、大容量内存、高IO吞吐、高数据交互可满足当前绝大多数生命科学计算需求。

6. 一站式咨询服务贯穿全业务周期

方案设计帮助用户“买好”、“建好”集群，专家团队保驾护航，帮助用户“用好”、“管好”集群。