

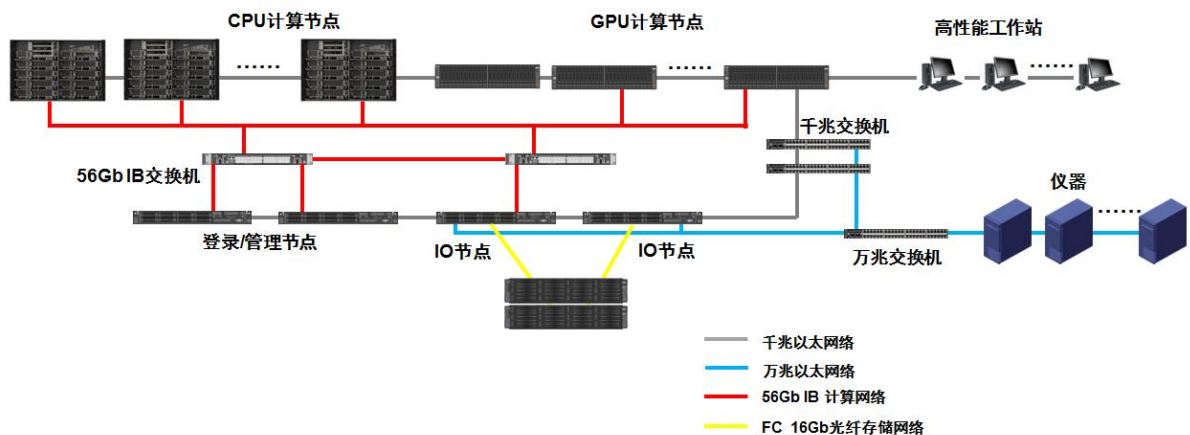
# 浙江大学冷冻电镜中心 高性能计算集群介绍

集群采用AMAX GPU 服务器和联想CPU服务器组合成一个高性能异构计算集群，满足整体架构要求。其中联想刀片服务器作为集群的CPU计算节点，AMAX机架式服务器作为集群的GPU计算节点。仪器产生的数据通过万兆交换机网络连接至IO节点，最终保存到后端存储。

用户通过管理登录节点提交作业，计算节点从IO节点读取后端存储数据进行并行计算，其结果将被保存至I/O节点，通过I/O节点放置到后端存储中。同时，配备3D效果显示工作站用于远程操作。

该异构集群资源调度使用Platform LSF作业调度系统，集群管理软件安装在管理节点上，用来负责整套系统的任务分配及各个节点的健康状态。

集群的计算网络使用高速的，无阻塞的56Gb/s infiniband网络，管理网络使用千兆以太网网络，整套系统使用双网独立的形式以避免管理网络与计算网络之间的冲突。详见如下拓扑图：



● 高性能计算平台分为 6 个子系统：计算子系统、管理登陆子系统、网络子系统、存储子系统、软件子系统及 3D 高性能工作站。其中，计算子系统包含 15 个 GPU 计算节点，38 个双路 CPU 计算节点，平台理论总峰值性能达到每秒 128.638 万亿次 (TFlops) 的浮点运算。存储子系统配置 648 TB NL SAS 磁盘。3D 高性能工作站共 12 台。

### **GPU 节点组成:**

10 台 AMAX 的高性能服务器 PSC-HB1X, 每个台配置 2 颗 Intel Xeon E5-2637v4 处理器、4 块 Nvidia GTX-1080TI, 11GB 显卡、256G DDR4 2400MHz 内存和 4 块 800G MLC SSD 硬盘;

3 台 AMAX 的高性能服务器 PSC-HB1X, 每个台配置 2 颗 Intel Xeon E5-2637v4 处理器、4 块 Nvidia Tesla P100, 16GB 显卡、512G DDR4 2400MHz 内存和 4 块 800G MLC SSD 硬盘;

2 台 AMAX 的高性能服务器 PSC-HB1X, 每个台配置 2 颗 Intel Xeon E5-2637v4 处理器、2 块 NVIDIA Quadro GP100, 16GB 显卡、512G DDR4 2400MHz 内存和 4 块 800G MLC SSD 硬盘;

1 台 IBM 高性能服务器 POWER, 每个节点配置 2 颗 8 核 3.259 GHz POWER8 处理器、4 块 Nvidia Tesla P100, 16GB 显卡、512G DDR4 2400MHz 内存和 2 块 480GB MLC SSD 硬盘。所有节点集成双口千兆网口及 56GB Infiniband HBA 卡。

### **CPU 节点组成:**

34 台联想 SD530 服务器, 每个节点配置 2 颗 Intel Xeon Gold 6140 (2.3GHz, 18 核 36 线程)、256GB DDR4 2666MHz ECC RDIMM 内存、1 块 300GB 12Gbps 10Krpm SAS 硬盘;

3 台联想 SR630 服务器, 每台配置 2 颗 Intel Xeon Gold 6140 (2.3GHz, 18 核 36 线程)、512GB DDR4 2666MHz ECC RDIMM 内存、1 块 300GB 12Gbps 10Krpm SAS 硬盘;

1 台联想 SR630 服务器, 每个节点配置 2 颗 Intel Xeon Gold 6128 (3.4GHz, 6 核 12 线程)、768GB DDR4 2666MHz ECC RDIMM 内存、1 块 300GB 12Gbps 10Krpm SAS 硬盘。每台服务器配置 2 个 1Gb 以太网口及 2 个 56Gb InfiniBand FDR 端口。

- **存储子系统:** 联想数谱 DS5760 集中存储, 配置 648 TB NL SAS 磁盘, 96GB 高速缓存, 4 个 16Gb FC 激活, 8 个 16G FC 接口, 设备主要包含 2 台 IO 节点和 1 台高速存储系统。
- **3D 高性能工作站:** 共 12 台, 每台配置 Intel I7-6700K 处理器、64G DDR4

2400MHz REG ECC DIMM 内存、Nvidia Quadro M4000 显卡、256G SSD+6T 企业级硬盘、27 英寸 2K 高分 3D 显示器、Nvidia 3D 眼镜套装。

- **网络子系统：**主要包含 1 台万兆以太网交换机、2 台千兆以太网交换机和 2 台 FDR InfiniBand 交换机。
- **集群软件子系统：**包含集群操作系统、集群管理软件、Platform LSF 作业调度软件、GPFS 并行文件系统和 Intel 集群工具集，还包含日常使用所需的标准工具、图形库。
- **管理子系统：**用于系统管理、作业调度和用户登录，用户通过前端 2 个管理登陆节点访问集群计算资源，作业调度系统安装在管理登陆节点，负责资源管理、作业调度、集群管理等管理，向最终用户和管理员提供统一使用、管理平台，包含管理服务器、资源管理软件、集群管理软件，以及搭建并行环境所需的软件。