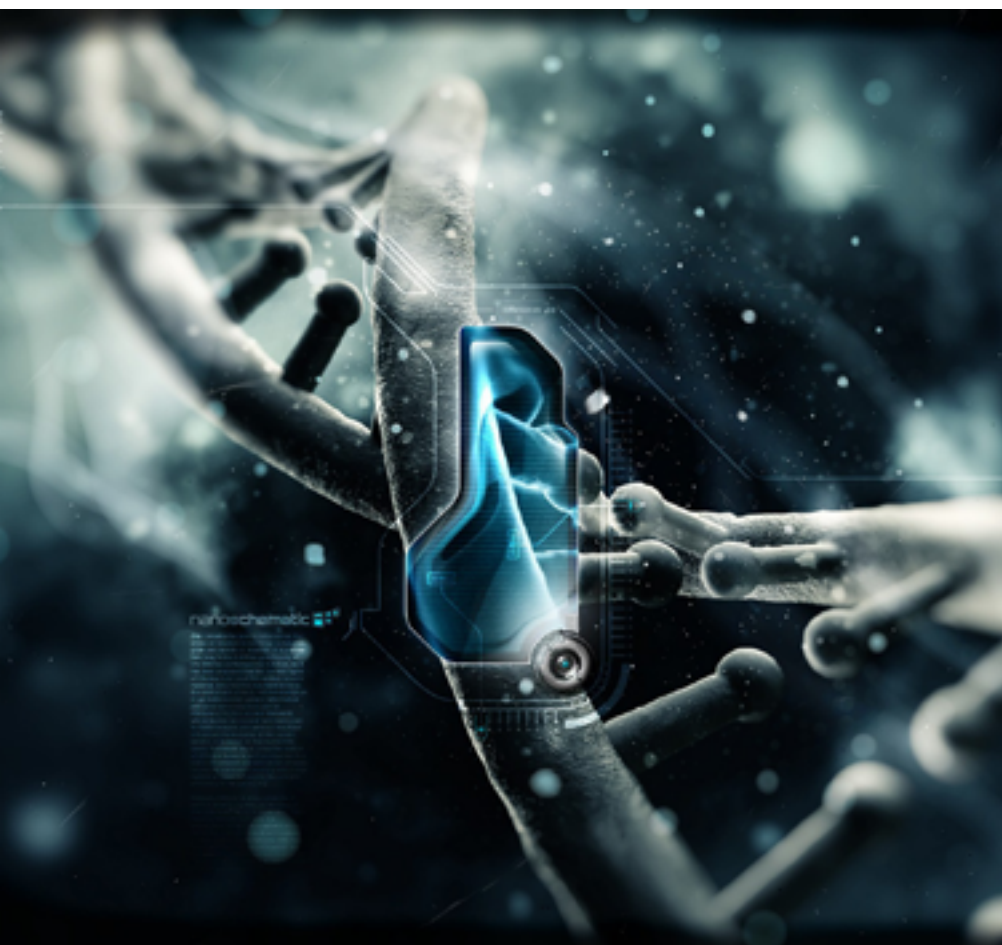


上海交大搭建大容量存储系统 保证 HPC 平台稳定可靠



客户资料



机构	上海交通大学高性能计算中心
行业	教育科研
国家	中国
网址	hpc.sjtu.edu.cn

高性能计算迎来新的发展机遇

高性能计算 (High Performance Computing , HPC) 被誉为 IT 行业 “ 金字塔上的明珠 ” 。它通过并行计算提高运算速度，以解决大规模科学计算和海量数据的处理问题。随着数据量以及数据价值的不断增长，HPC 的应用领域不断拓展，在科学计算、天文、医学、能源、生命科学以及仿真设计等领域发挥着重要的作用。据 IDC 预测，到 2019 年，全球 HPC 市场将达到 152 亿美元的规模。

随着 HPC 应用领域和范围不断扩大，其负载程度与复杂度都在不断增加，数据量更是爆发式增长，这对计算性能和存储性能都提出了更高的要求，并不断挑战技术的极限。随着中国 “ 十三五 ” 的到来，各行各业向数字化转型已成为大趋势，这无疑将对 HPC 的应用发展起到巨大的推动作用，并帮助提升各行业的竞争力。如何充分发挥 HPC 的独特价值，使其应用领域更加拓宽、应用更加深入和大众化，促进各行业更快向数字化转型，并在数字经济中提升竞争力，已经成为未来的发展方向及关注点。

在 HPC 的实际应用中，计算、存储、网络、软件等各方面不同程度的 “ 短板 ” 都在影响系统的整体计算能力。如何将各部分的性能相协调并实现创新，是推动 HPC 快速发展和广泛应用的关键。在 HPC 中，高效地将数据传入和传出计算节点至关重要，这通常会涉及到一些复杂的因素。与越来越强大的 HPC 计算处理能力相比，存储组件往往会成为整个 HPC 系统的瓶颈。数据量的快速增长，常常使存储容量相形见绌，而日益复杂的存储管理和运维，也加重了存储管理员的负担。另外，存储系统

“ 我们上海交大在理工、医学方面的学科是强项。这方面的科研和教学课题组有大量建模、仿真或科学计算需要用到高性能计算，这些学科对于数据的要求比较高，同时数据量也很大，对存储的需求增长很快，而我们现有的 HPC 存储系统不给力。因此，我们在听取了大量用户的反馈意见后，决定对 HPC 存储系统进行升级。戴尔的 HPC 存储方案不仅完美地解决了我们遇到的困扰，同时，戴尔提供的完善的服务帮助我们快速完成了项目的安装、测试和调优工作。这套系统上线后的表现令人非常满意。 ”

上海交通大学高性能计算中心韦建文

难以扩展也是 HPC 面临的挑战之一。数据对性能和容量的需求持续增加，系统管理员除了需要进行精心的规划和充分的配置外，还亟需能够弹性扩展计算和存储资源，从而为整个 HPC 系统提供有力的支持。

随着 HPC 的快速发展，高校作为科学研究的重要机构，其教师、学生以及科研人员对 HPC 的需求也在不断增加和提升。近年来，许多高校开始兴建自己的 HPC 中心，并有意识地把小的散布在院系中的计算集群整合为校级的集中式的 HPC 中心。中桥调研咨询 2016 年年初对教育行业信息主管进行的调查显示，有 34.1% 的受访者表示，在未来 24 个月，其所在高校或教学研究机构将部署 HPC。可以预见，未来 HPC 将在高校和教育研究机构迅速发展，对高校的教育、科研等将起到越来越重要的作用。

上海交大高性能计算面临挑战

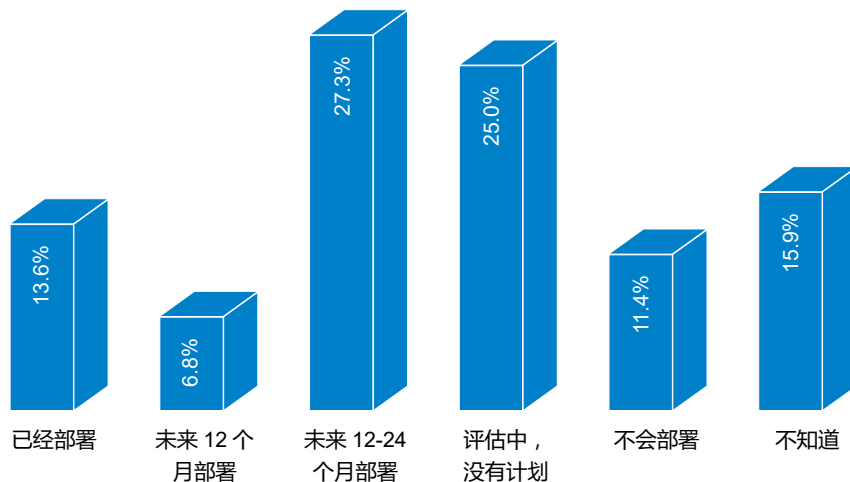
上海交通大学（以下简称“上海交大”）是一所涵盖理、工、医、经、管、文、法等 9 大学科门类的综合性全国重点大学。作为中国著名的高等学府，该校在工学、商学、医学领域拥有很强的学术影响力，其对高性能计算的建设和应用在国内高校居领先水平，并推动国内高校对高性能计算中心发展新模式的探索。

上海交大高性能计算中心始建于 2013 年，同年 4 月完成整机安装并试运行，10 月正式对外服务。上海交大 HPC 系统重点支持校内各院系的教学和科研，为不同的科研和教学课题组服务，目前是全国高校中使用异构计算比例最高的 HPC 中心，共有 332 个 CPU 节点、130 块 GPU 加速卡和 65 个 GPU 加速卡节点。截至目前，他们已经累计为 130 多个课题组提供了服务，每年的计算量超过 2300 万核小时 CPU，系统的整体资源利用率保持在 60% 以上，而 CPU 的利用率则常年超过 70%。

不过，伴随着用户数量的增长和所支持的学科课题组的增多，运算量也日益飙升，这给从成立之后系统一直没有升级的 HPC 中心带来了不小的压力，尤其在存储方面，所面对的挑战更多，迫切需要通过系统的升级来应对。

你所在的高校或教育研究机构是否考虑部署 HPC ?

(所有受访者, N=44)



数据来源：中桥调研咨询关于“十三五”IT 新趋势和新动向的调查，2016 年 5 月

挑战之一：存储容量亟待扩容

首先摆在上海交大 HPC 中心面前的问题是存储容量不能满足需求。由于理工科是上海交大的强项，比如，物理、天文、生命科学等自然科学以及一些新兴学科都是使用 HPC 较多的学科，而这些学科的课题研究本身对于数据的要求较高，所产生的数据量也较大。从存储容量上看，上海交大 HPC 中心建成时，系统的可用容量 500 多 TB，到 2016 年年初时，存储的利用率已经超过 90%，而数据量还在不断攀升，不但使存储容量告急，也导致了存储性能不够稳定。因此，系统急需扩大存储容量，并提升访问速度，以保障用户的正常使用和良好的体验。

挑战之二：难以较好地支持小文件读写操作

在上海交大 HPC 中心，许多学科的大规模并行计算都需要小文件操作，如基因测序等应用。因此，对存储系统的响应速度要求较高，否则，当系统产生的数据量激增，即便计算系统速度再快，如果并行文件系统无法快速响应，那么也会急剧降低 HPC 系统的整体性能。上海交大 HPC 的存储系统性能不稳定，尤其是针对小文件读写速度较慢，严重影响了用户体验的流畅性，需要对容量和性能、读写速度有质的提升。

挑战之三：系统可管理性差

要让 HPC 集群发挥作用，就需要高效、长期的运维机制。但是从中国现阶段的 HPC 运维工作来看，大多数 HPC 中心缺乏专业的运维团队，这就导致高性能计算机的故障多、性能低、寿命短，无法真正发挥高性能计算机的作用。上海交大 HPC 中心在这方面也遇到了这样的难题。由于 HPC 系统建成仅三年时间，针对 HPC 运维经验相对不足，尤其是前期使用的存储系统采用在自命令行下操作文件系统，不直观，可管理性差，更加大了运维工作的难度。



HPC 存储系统升级需求

上海交大 HPC 中心针对 HPC 存储系统面临的诸多挑战，决定对该系统进行升级。通过升级存储系统，扩大存储容量、提升存储性能，同时改善系统的可管理性。具体需求如下：

需求一：扩容

存储是决定 HPC 系统性能、服务水准和用户体验等的关键因素之一。针对上海交大 HPC 系统正面临的一些问题，如何高效解决伴随用户数量和数据量增长而带来的存储容量不足问题，经济高效地实现存储容量的扩展升级，是上海交大在选择新系统时的第一诉求。

需求二：优化对小文件的读写

对上海交大 HPC 用户来说，在使用 HPC 进行课题研究的过程中经常涉及到小文件读写操作。对小文件读写速度的提升，可以从整体上保证存储系统的稳定，从而提升计算速度。

需求三：提升可管理性

在 IT 预算有限的情况下，存储系统的升级除了要能解决上海交大 HPC 系统现阶段所遇到的容量和性能问题外，还要能提升系统的可管理性，降低后续的系统运维成本。

需求四：完善的服务支持

上海交大 HPC 中心在选择存储方案提供商时，还需要考虑方案提供商在 HPC 领域的经验积累，以及提供整体解决方案的能力和完善的本地售后服务支持，降低项目实施中的风险，加速 IT 服务交付管理效率。

上海交大 HPC 存储系统方案选择

上海交大 HPC 中心在选择存储系统方案时，经过多方甄选和测试，最终选择了戴尔公司提供的解决方案——安装了 Intel EE for Lustre 软件的戴尔 HPC 存储解决方案。上海交大 HPC 中心选择戴尔解决方案主要基于以下理由：

- **高性价比。** 上海交大 HPC 中心在采购时研究对比了多个厂商提供的方案，在价格与容量、吞吐性能的比较中，戴尔提供了一个性价比很高的解决方案。
- **得以验证的参考架构。** 戴尔的解决方案中，除了存储设备具有易于使用、大容量、高吞吐量、横向扩展，并提供较低的 TCO 之外，戴尔与英特尔有紧密的合作关系，其存储系统采用英特尔的 Lustre 文件系统，整套系统通过了英特尔的验证。该套系统开放、透明，其技术指标、测试实施方法、

测试软硬件配置都非常透明，给用户以信心，所提供的参考架构已在全球许多案例中得到了验证。

- **完善的售后服务。** 戴尔提供了完善的本地售后服务支持，包括工程师上门、设备替换、备件及时送货等服务。
- **众多成功案例。** 戴尔的存储系统在 HPC 领域有许多成功的案例，甚至在

国外高校也有与上海交大类似的应用。另外，戴尔在数据中心领域占有率较高，上海交大之前采购过戴尔数据中心其他产品，对其售后服务、产品性能和可靠性非常满意。

经过安装测试，新的存储系统目前已经作为主存储系统在上海交大 HPC 中心正常运行。之前所有的应用和数据已经全部迁移至新的系统上，并完全满足了 HPC 系统对存储的需求。

上海交大 HPC 系统存储 IT 环境设备构成：

- 6 台戴尔 PowerEdge R730 机架服务器，作为元数据服务器和对象数据服务器
- 1 台戴尔 PowerVault MD3420 磁盘阵列，作为元数据存储系统
- 1 台戴尔 PowerVault MD3460 磁盘阵列，作为对象数据存储系统
- Extreme X440 以太网交换机、IB 交换机扩展卡和线缆
- 英特尔 Lustre 并行文件系统



上海交大 HPC 存储系统解决方案的优势

存储

戴尔 PowerVault MD3 产品包括一系列高性能 SAN、直接连接的存储阵列和高容量密度阵列（均使用扩展盘柜）。本项目中使用了 MD3460 和 MD3420。

- 消除容量瓶颈。戴尔 PowerVault MD3 高密度阵列最高可存储 540TB 数据，向上扩容至 180 块硬盘。其模块化架构可简便地实现扩容，为未来业务发展提供保障。
- 保证性能稳定。戴尔 MD3460 和 MD3420 通过双 12G SAS 控制器提供高速连接，并且每个阵列的高速缓存容量达到 16GB，保证上海交大对 HPC 系统容量和性能的需求。此外，戴尔 PowerVault MD3460 自带高性能层许可，通过自动分层处理满足对象存储管理数据的性能要求。
- 经济高效。戴尔 PowerVault MD3 通过模块化架构，实现按需升级，节省存储成本，同时降低了能耗、制冷和空间。此外，在基本的产品架构上还提供了企业级的功能，提高了操作的灵活性。

服务器

- 提供高性能。PowerEdge R730 服务器采用英特尔至强 E5-2600 V3 系列处理器，配有 7 个 PCIe3.0 扩展插槽，内存容量可从 64GB 可扩展到 128GB，可以最大限度地提高各种类型工作负载的计算性能；通过 16 个 12Gbps SAS 驱动器和高性能双 RAID，加快应用程序的数据访问速度；利用高级加速器和 GPU，能够在 HPC 环境中最大限度地提高性能。
- 提高运营效益。PowerEdge 服务器自动化的部署流程有助于减少手动步骤，并降低出错的可能性；iDRAC 提高服务器远程管理效率和生命周期管理效率，进而提高 IT 效率。

管理运维

该项目所采用的英特尔文件系统 Intel Manager for Lustre (IML) 具有智能、丰富且直观的管理界面，大大简化了所有硬件和存储系统组件的管理和监控。它在容量和性能方面易于扩展，从而为未来增长提供了方便的途径。它通过提供一个用于进行管理的集中式 Web GUI，简化 Lustre 文件系统的管理，直观地进行监控和操作，如启动从一个节点到另一个节点（对于 OSS 或 MDS）的文件系统故障转移、格式化文件系统、发出对目标的装入和卸载命令、监控 Lustre 文件系统的性能和其各个组件的状态。

用户收益

通过安装戴尔 HPC 存储系统解决方案，上海交大 HPC 中心成功解决了该中心 HPC 系统的痛点，不仅应对了用户日益增长的存储要求，提升了存储系统性能的稳定性和提供了轻松、高效、安全的管理环境。这不仅有利于满足数学科研等相关课题组的需求，强化了校内各院系科研实力，保证各院系课题组科研和教学的正常、顺利进行，更有助于提升上海交大科研和教学的竞争力。

突破存储容量和性能瓶颈

新系统上线后，目前投入使用的存储容量达到了 2PB，几乎是原有存储系统容量的 3 倍，后续还会继续扩容至 3PB。文件系统的负载目前可保持在一个较低的水平，同时小文件的读写操作速度很快。

生命科学学院基因组测序课题组一直是 HPC 系统的用户。新系统上线后，他们不再受原来系统在容量或性能上的限制，已经在新的系统上新增存储 100TB。

课题组的研究人员说：“与以前的系统相比，新系统运行更加稳定，没有出现什么问题导致整个集群性能下降，可以同时利用 100 多个节点进行高通量的基因测序分析。”

简化系统管理难度

英特尔文件系统 Lustre 智能、丰富且直观的管理界面大大简化了所有硬件和存储系统组件的管理和监控。

上海交大 HPC 中心韦建文老师表示：“在可管理性上，英特尔 Lustre 提供了一个 Web 界面，使我们能够非常方便地查看整个文件系统的吞吐量和故障点，并且当出现故障时操作也非常简单和方便。管理员只要远程登录到管理界面，鼠标点击几下就能完成故障恢复，比以前要方便很多。”

稳定的文件系统性能

在保障文件系统性能稳定方面，英特尔 Lustre 并行文件系统搭配戴尔 HPC 存储设备这一得以验证的参考架构，对于 Lustre 并行文件系统有针对性地进行了优化，提高了文件读取速度。

上海交大 HPC 中心韦建文老师说：“我们对于系统的小文件操作非常满意。在我们系统运行过程中，整个文件系统的负载可以保持在非常低的水平。”

如欲浏览戴尔的所有案例，请登录 Dell.com/casestudies

