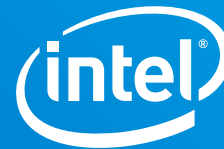


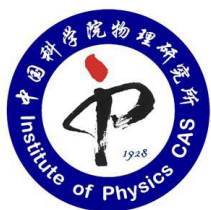
案例研究

第二代英特尔® 至强® 可扩展处理器
英特尔® Omni-Path 架构
高通量计算
材料基因组计划



中科院借力数据增强型超算技术 加速材料基因探索

英特尔与戴尔携手，助中国科学院物理研究所进一步优化基于高通量计算技术的材料基因计算与数据处理平台的算力与存储能力，以提升新材料研究与成果共享的效率



“积极响应高性能计算领域的全新技术和应用趋势，充分利用高通量计算来加速材料基因计算，以提升新材料研发效率，是当今材料研究领域的重要选择之一。通过采用第二代英特尔® 至强® 可扩展处理器、英特尔® OPA 以及戴尔易安信 PowerEdge 服务器、PowerVault ME4 系列和 Isilon 系列存储等先进产品与方案，我们的新平台在性能、可用性、可靠性以及易维护性等方面都收获了一线材料研究人员的好评，让他们在新材料研究领域获得了更强的助力。”

刘淼
特聘研究员
中国科学院物理研究所

传统的材料研究方法因其研发时间长，成功率不确定，正渐渐成为制约工业产品创新和质量优化的一大瓶颈。为加速新材料研究步伐，强化国家工业竞争优势，中国科学院物理研究所（以下简称“中科院物理所”）开始更多地借助高性能计算领域近年来涌现的全新技术分支——高通量计算，通过建立高通量材料基因计算和数据处理平台来支持材料基因组项目的推进，并基于材料基因数据库与云资源平台为全国范围内的更多材料研究人员提供数据共享服务。

与更专注于提升速度、着重增强单个计算任务的效率，最终以“算得快”为关键指标的传统高性能计算不同，新兴的数据增强型高性能计算技术——高通量计算更看重“算得多”这一能力，也就是在提升工作任务并行度，及其处理海量数据时的吞吐量上下苦功。基于此，高通量计算平台往往在并行处理性能、平均响应时间、吞吐率、扩展性以及成本等方面有着更高，或更为严苛的要求。

从这些现实需求出发，中科院物理所与致力于输出核心算力创新的英特尔，以及提供服务器、存储设备及解决方案的戴尔携手，通过导入第二代英特尔® 至强® 可扩展处理器、英特尔® Omni-Path 架构（英特尔® OPA）等关键计算与高速互连技术，以及基于这些产品技术，并融入戴尔独有创新功能和特性的戴尔易安信 PowerEdge 服务器、戴尔易安信 PowerVault ME4 系列和 Isilon 系列存储设备与方案，打造出了面向未来研究需求的新一期材料基因计算和数据处理平台（以下亦简称为新平台），不仅达成了提升材料基因计算效率的主要目标，也兼顾了为广大材料研究人员提供完整、齐备的共享材料基因数据库的需求。

中科院物理所获得的应用优势：

- 第二代英特尔® 至强® 可扩展处理器的部署，帮助中科院物理所以适宜的线程成本，获得了更为出色的高通量材料基因算力，使其有信心应对每年数以万计的材料计算任务，大大加快中国的新材料研究步伐；
- 来自戴尔的多款久经考验的计算、存储产品及领先的超融合设备，与第二代英特尔® 至强® 可扩展处理器和英特尔® OPA 相结合，可帮助新平台实现高通量材料基因计算中的高性能数据闭环，不仅能助力中科院物理所构建起世界一流的材料基因数据库，更能帮助其演进为中国材料基因研究的“动力源”。

如果说石油是现代工业的“血液”，那么材料扮演的，就是其“筋骨肌肉”的角色。工业产品要想具备更强的竞争力，持续地“脱胎换骨”，即坚持研发和导入新型材料就势在必行。此前，材料领域的研究人员主要是通过不断地实验试错，研究不同材料的表征结构来探索新材料的性质。随着现代材料的组成越来越复杂，尤其是各类复合材料的出现，使得很多情况下的常规实验手段已难以奏效，其开发周期长、开发成本高且成功率较低等问题也逐渐显现。

随着材料基因组计划 (Materials Genome Initiative, MGI)¹ 在全球范围内的启动，以及密度泛函理论 (Density Functional Theory, DFT)、蒙特卡洛方法 (Monte Carlo Method) 等理论模拟方法的日益成熟，材料领域的研究人员迎来了破局点，他们开始更多地利用计算机方法，尤其是高通量计算方法提供的强大并行算力，以模拟计算的方式来进行材料筛选等工作。

以目前备受瞩目的半导体材料研究为例，过去该领域每一种新材料的运用和表现，例如薄膜在衬底上的生长行为，是否钝化或出现表面重构，都需要进行反复的测试和验证。如今研究人员可利用 DFT 理论，通过部分已知的密度泛函计算结果，对体系的总能量进行拟合，以量子力学方法来精确计算材料原子间的电子相互作用，从而获得更多化合物的电子结构、磁性和结构稳定性数据，进而从理论上指导材料设计，加速更多新材料在半导体中的运用。

作为中国首屈一指的材料科学研究与应用机构，中科院物理所也紧随潮流，致力于运用高通量计算方法开展新材料研究，并已取得丰硕成果，在超导体、纳米以及动力电池等材料方向上获得了诸多突破。在中科院物理所看来，加速中国新材料研究步伐的重要任务之一，就是要坚持将信息技术的最新发展融合到材料研究中去，以更强、更高效的计算力，构建完备的材料基因数据平台，帮助研究人员省去繁琐的计算环境采购、搭建工作，让他们能将更多精力聚焦于材料研究本身。

为此，中科院物理所与英特尔和戴尔携手开展技术合作，引入领先的计算、互连技术及成熟的数据处理、存储产品和方案，为新一期高通量材料计算和数据处理平台的建设提供强有力的 IT 基础设施支持，致力于将之打造为中国材料基因计算、同时也是中国材料基因研究领域的强力引擎。

解构高通量材料基因计算解决方案

如前文所述，在中科院物理所新平台中扮演关键角色的技术是高通量计算，它可以被简单理解为能够并行处理大量独立工作

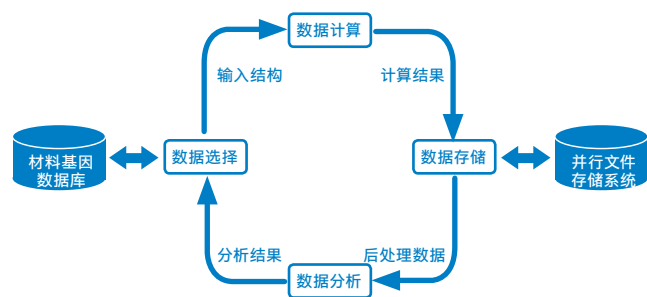
任务的计算系统，与传统的高性能计算系统相比，其特点在于：

- **工作特性：**通常采用线程级并行处理，工作负载根据任务请求不断变化；
- **执行效率：**更强调计算与数据的集成，数据存储读写性能直接影响系统效率；
- **性能目标：**更注重提升单位时间内，并发的计算与数据处理需求数量；
- **成本要求：**系统处理能力与线程数量正相关，单线程成本优先可提升整体处理能力。

总的来说，高通量计算能“大批量”处理工作负载的特性，非常适用于需要进行大量样品筛选的场景，如生物、医药等领域的研发过程。在材料基因研究过程中，同样也可以采用高通量计算方法来进行海量的材料基因模拟计算和筛选。

如图一所示，材料基因计算通常可简化为以下步骤：

- ① 系统从外部的材料基因数据库中选择数据，产生可被特定的模拟计算软件 (例如 VASP²) 调用的文件；
- ② 通过高性能的并行计算，系统获得相应的材料性质数据，例如材料的能量密度、电子结构、合成时的能量增减量等；
- ③ 系统将计算结果保存至内部存储系统以便进行下一步分析，通常需采用高性能的并行分布式文件系统；
- ④ 数据经后处理整合后，可开展进一步深度分析，并使分析结果用于材料筛选等，同时分析结果也可以导入材料基因数据库，由此形成完善的程序化计算流程。



图一 材料基因计算的工作流程³

上述流程涉及的结构文件生成、理论模拟计算、计算结果分析等流程，都可以通过一定方式形成一系列相互独立的计算指令

与工作任务的，并交由特定的软件包予以执行，这就为高通量计算方法的运用提供了合适的场景。

在执行高通量材料基因计算任务的同时，中科院物理所还计划在新平台中构建一个数据中心，来更为高效地存储计算所得的数据并进行后续分析。同时它还计划以私有云的方式提供一个共享数据平台，将材料基因计算过程中的数据、代码以及计算工具等予以共享，以帮助全国范围内更多材料研究人员节省因购买同类设备、搭建类似环境所耗费的人力和物力成本，让他们能更加专注于自己所从事的研究工作。

基于这一整体构想，中科院物理所计划通过四个关键软件包，来部署其高通量材料基因计算解决方案：首先是基于 SLURM (Simple Linux Utility for Resource Management) 资源管理系统的高通量作业调度包，其主要对各个计算分析作业进行调度、监控，确保执行效率最优化；其次是数据归档包，用于将计算结果在 Lustre 并行分布式文件系统中高速归档；接下来的数据后处理包主要负责将前面的原始计算结果进行推行和整合，供后续数据分析使用或进行图形化展示/查询；最后是大数据分析包，其部署后将帮助研究人员基于海量的计算数据和分析结果，通过人工智能 (Artificial Intelligence, AI) 等方式，更高效地开展新材料筛选、合成效果预测等工作。

优秀的设计思路需要适宜的硬件基础设施来予以支撑，而这对于更专注于材料研究工作，对 IT 技术缺乏专业优势和积累的中科院物理所而言，也是一个不小的挑战。为了让新平台更充分地释放潜能，在未来的计算任务中发挥更大功效，中科院物理所本着“把专业的事交给专业的人去做”的出发点，引入了英特尔与戴尔作为创新伙伴，一同基于以上软件层面的完整框架，量身定制高性能、高可用的计算与存储硬件架构。在三方开始行动之前，中科院物理所也针对新平台的基础设施提出了几点要求：

- **技术先进性：**要求新平台的计算与存储硬件架构要能支撑所内未来数年内持续增长的高通量材料基因计算任务，并保持一定的性能优势；
- **工作稳定性：**新平台建设完成后，需肩负起海量的高通量材料基因计算任务，而任何系统故障都会带来难以预知的损失，因此架构需要更加注重可靠性和稳定性；
- **运维便捷性：**中科院物理所相对紧张的人力资源，使其没有更多余力为平台配备专职的运维专家，因此需要计算与存储硬件架构在运维上做到便捷易上手。

英特尔® 至强® 加量不加价 带来更优并行算力

面对中科院物理所提出的这几个要求，首当其冲的，就是实现新平台的技术先进性，换言之，也就是要满足其在算力，或者说性能上的要求。

谈到新平台的算力，它的主要输出源就是处理器，因此针对中科院物理所的高通量材料基因计算需求来选择一款合适的主力处理器就是整个平台构建的重中之重，它需要多维度的、全面且均衡的考量。具体来说，由于高通量计算方法更看重并行处理能力，要求计算系统能在同一时刻处理尽可能多的任务，有着与线程数量呈正相关的特点，因此处理器的核心和线程数是首要考量因素；其次，由于处理器的各个核心和线程均共享 L3，或者说是末级缓存，因此更大的 L3 缓存容量也是要列入考虑的，它能保证每个并行任务分配到更多的缓存，提高缓存命中率；第三，处理器频率的提升也不可或缺，它有助于提升单个工作任务的处理速度；最后，价格或成本也是必须要考虑的因素，虽然更高配置的处理器可以在上述规格和指标上有更好的表现，但也会抬升成本，相对而言，更合适的选择应该是在保证并行处理性能的前提下，实现 CAPEX (Capital Expenditure, 投资成本) 获益的最大化，所以需遵循单线程成本优先的原则。

为帮助中科院物理所选择更具性价比的处理器，英特尔与所内专家一起，对旗下多款型号的处理器进行了全面测试和评估。最初中科院物理所看中的是英特尔® 至强® 金牌 6230 处理器，但随着新平台未来需承载更多计算进程的规划出炉，中科院物理所又打算引入主频相对更高的英特尔® 至强® 金牌 6248 处理器。而英特尔在帮助中科院物理所进行了全方位的测试后给出了如下建议：虽然至强® 金牌 6248 处理器凭借更高的主频带来了性能的部分强化，但由于它与至强® 金牌 6230 处理器的核心和线程数相同，因此这一增强还谈不上有根本上的突破。在认真评估了第二代英特尔® 至强® 可扩展处理器产品家族于 2020 年年初发布的升级型号之后，英特尔建议中科院物理所选用其中的英特尔® 至强® 金牌 6230R 处理器。

与 2019 年发布的第二代英特尔® 至强® 可扩展处理器相比，2020 年初发布的几款升级型号在睿频、核心和线程数量以及缓存容量上都进行了强化。就以英特尔® 至强® 金牌 6230R 处理器为例，如表一所示，其核心和线程数量、以及缓存容量与升级前的型号 (英特尔® 至强® 金牌 6230 处理器) 相比均提升了 30%，同时睿频频率也从 3.9GHz 提升到了 4.0GHz，但成本基本持平⁴，这意味着这一选择可让新平台的 CAPEX 收益获得明显的提升。

	主频	睿频	核心数	线程数	缓存
英特尔® 至强® 金牌 6230R 处理器	2.10 GHz	4.00 GHz ↑	26 ↑	52 ↑	35.75 MB ↑
英特尔® 至强® 金牌 6230 处理器	2.10 GHz	3.90 GHz	20	40	27.5 MB

表一 英特尔® 至强® 金牌 6230R 处理器 (2020 年) 与英特尔® 至强® 金牌 6230 处理器 (2019 年) 在主要产品规格上的对比

同时, 为帮助中科院物理所应对某些特定场景对计算性能的更高需求, 英特尔也为其推荐了英特尔® 至强® 铂金 9242 处理器。这款处理器集成了 48 个核心, 支持 96 个线程的高并发, 并支持 12 通道的 DDR4 内存, 能够以更强悍的算力, 支持中科院物理所在材料基因算法优化、基于机器学习及深度学习的材料特性预测等应用上开展更多探索。



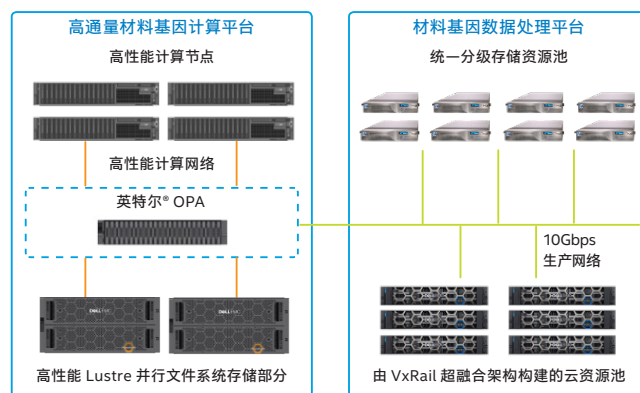
被誉为“现代化数据中心基石”的**戴尔易安信 PowerEdge 服务器**是承载中科院物理所新一期高通量材料计算和数据处理平台算力需求的关键设备。它可搭载第二代英特尔® 至强® 可扩展处理器, 并凭借高度可扩展的体系结构和灵活的内部存储, 为广泛多样的应用程序提供更高性能支持。它“预装 Vmware 虚拟化软件, 自带云基因”, “高可扩展性的业务体系结构”, “智能自动化, 易于远程运维”等特点, 非常有助于加速企业的 IT 转型和业务创新⁵。

戴尔存储产品搭配英特尔® OPA 实现计算与数据紧密集成

解决了新平台在算力上的特定需求后, 中科院物理所与英特尔和戴尔开始考虑强化新平台上计算与数据的集成, 这一点正是高通量计算技术在“数据增强型”上的具体体现, 也是它在并行算力之外另一个较为关注的性能要素, 它对存储模块的吞吐量、时延和带宽性能有更高要求。

为此, 英特尔与戴尔一起, 将戴尔易安信 PowerVault ME4 系列存储、戴尔易安信 Isilon 系列可扩展 NAS 存储、戴尔易安信 VxRail 超融合架构以及英特尔® OPA 等先进软硬件产品与技术引入新平台。

如图二所示, 根据中科院物理所在材料基因计算、数据存储、数据分析以及展示等方面的需求, 新平台的主要硬件架构会分为高通量材料基因计算和材料基因数据处理两个平台, 前者主要承载高通量计算任务以及计算过程中的高速文件存储, 而后者则提供了数据存储、分析以及云化共享能力, 两个平台间以 10Gbps 生产网络相互连接。



图二 中科院物理所新一期高通量材料计算和数据处理平台主要硬件架构⁶

在高通量材料基因计算的运行过程中, 成百上千的并行独立计算任务会产生大量的过程文件, 这些文件需要被高速存储起来以备执行后续的数据推行和整合。为此, 新平台采用了高性能的 Lustre 并行文件系统, 并由两台基于英特尔® 架构处理器的戴尔易安信 PowerVault ME4084 高密度存储设备 (2*84*8TB, 总容量共计 1,344TB) 来对其进行承载。Lustre 文件系统不仅可支持数百 PB 数据存储空间以及数 TB 每秒的并发带宽, 还具有良好的扩展性, 中科院物理所可以灵活便利地根据需求对 ME4 系列存储设备进行扩展。



基于戴尔先进架构构建的**戴尔易安信 PowerVault ME4 系列存储** 配备有英特尔® 至强® 处理器, 针对 SAN/DAS 进行了优化, 可帮助企业用户快速建设及整合便捷、高性能和经济实惠的存储系统。ME4 系列功能广泛, 支持各类驱动器类型、多种协议和全包式软件功能, 并提供大量扩展存储模块选项, 非常适于加速高性能计算、备份、VDI 等企业级应用程序的性能⁷。

材料基因计算得到的结果是大量非结构化数据，以电子结构(电子云密度)为例，单个材料的电子云密度文件就可达 10GB，且以图像形式存在。因此新平台需要为大量数据结果提供高性能、可扩展的存储系统。针对这一需求，戴尔为之提供了以出色的效率和卓越的可扩展性而著称的戴尔易安信 Isilon 系列可扩展 NAS 存储设备，并采用 4 台 H400 Isilon 混合横向扩展 NAS 机和 8 台 A2000 Isilon 归档横向扩展 NAS 机构成统一的分级存储资源池，分别承载数据分析和数据安全存储的重任。这两款产品不仅具备出色的横向扩展特性，可帮助平台弹性地扩展存储容量，其内置的 OneFS 操作系统更可提供高达 80% 的存储利用率，并具备多种数据保护和安全性措施来确保平台数据的安全性⁹，而这也满足了中科院物理所对新平台基础设施稳定性和可靠性方面的需求。



戴尔易安信 Isilon 系列可扩展 NAS 存储由 OneFS 操作系统提供支持，是一个功能多样的混合 NAS 阵列，可在性能、容量和价值之间实现更优的均衡。通过内置的多协议，它为用户提供灵活、便捷的可操作性，可帮助企业在单个平台上支持广泛的非结构化数据工作负载，以达到整合文件数据、消除成本高昂的存储孤岛、简化管理的功效⁹。

同时，戴尔还为新平台提供了 6 台戴尔易安信 VxRail 超融合一体机，来构建用于资源共享的云资源池。如前文所述，中科院物理所建设新平台的目标，也包括了将材料基因计算中的数据结果、代码和计算工具进行共享，以帮助更多材料领域的研究人员提升其研究效率。VxRail 超融合架构能够与戴尔的 SDDC (Software Defined Data Center, 软件定义数据中心) 软件体系完全集成，使平台能方便地基于 VxRail 超融合一体机部署 VMware Cloud Foundation 云环境。这一全面整合的平台带来的优势就在于：能大大简化私有云从规划、搭建到部署的复杂过程，并大幅降低后续的运维复杂度，使中科院物理所获得“一键上云”的能力，并如其最初要求的那样，尽可能为新平台实现更高的运维便捷性。



作为与 VMware 联合工程设计的超融合平台，**戴尔易安信 VxRail 超融合一体机**通过与 VMware 的深度集成，为用户提供了一键部署、一键升级、开箱既用的云平台交付能力，并实现了从硬件到软件，直至云平台的全生命周期管理，从而带来了云平台运维及管理上的充分便利，可帮助用户从小到大，按需扩展，轻松实现灵活的资源调配¹⁰。

为了让存储设备与计算节点更高效地协同工作，新平台还采用了英特尔® OPA 来提升互连效率和扩展性。英特尔® OPA 不仅通过链路层的技术革新优化了数据流控制，增强了数据包完整性保护，使延迟得到进一步降低，更为重要的是，基于英特尔® OPA 的交换机芯片可支持 48 端口，这帮助新平台仅部署单台交换机就可支持上千个端口(多达 1152 个端口)，来实现出色的可扩展性，以满足中科院物理所未来对新平台进行进一步升级和扩展的需求。

成果与展望

通过与英特尔和戴尔一同推进的三方协作，中科院物理所目前已在其中高通量材料基因计算新平台上完成了 160 个英特尔® 至强® 金牌 6230R 处理器节点、20 个高性能英特尔® 至强® 铂金 9242 处理器节点的部署，并在材料基因数据处理平台所在的数据中心中构建了完整的统一分级存储资源池，以及用于资源共享的数据云平台。

上述节点全部投入使用后，预计每天可执行 3,000 个基于 DFT 理论的作业任务，完成 100-700 个(视材料复杂度不同)无机晶体材料的基因计算，全年即可完成目前人类已知的所有无机晶体材料(去重后约 10 万个)的计算。而在此后的三到五年内，中科院物理所还将利用新平台提供的上述计算能力，对人类未知材料世界继续展开探索，将 50 万个未知材料的预测数据纳入平台的材料数据库中¹¹，从而将材料研究方法由传统的“发现-计算”转变为“计算-发现”，使信息技术真正转变为新材料研究的强大驱动力。

在加速材料领域研究之余，新平台的建设也将为“产-学-研”的结合提供更好的载体。通过共享云资源池，来自高校、科研机构以及企业的师生和研究人员可以利用数据平台中已有的材料基因数据快速开展下一步的新材料合成，有效避免了因重复计算带来的资源浪费。更值得一提的是，利用数据平台积累的丰沛数据，中科院物理所还可通过各类可视化方式向高校师生立体生动地演示各类无机晶体材料的内部结构，甚至可通过引入虚拟现实 (Virtual Reality, VR) 技术，让材料的学习和研究变得“触手可及”。

着眼未来，中科院物理所还计划与英特尔和戴尔等 IT 行业合作伙伴一起，继续优化和扩展材料基因计算平台规模，使高通量计算节点数量达到 500 个左右，算力提升达目前的两倍以上¹²，以跻身世界级顶尖材料基因数据平台。同时，新平台的计算对象也将由现在的无机晶体材料扩展到分子材料计算，令新材料研究范围变得更为宽广。英特尔和戴尔也将发挥自身优势，以性能和功能上持续更新换代的产品、技术和解决方案，助力中科院物理所在材料基因计算领域获取更多科研成果。

解决方案提供商



¹ 更多材料基因组计划信息，请参阅：<http://www.mgi.gov>

² VASP 软件全称 Vienna Ab-initio Simulation Package，是由维也纳大学 Hafner 小组所开发的电子结构计算和量子力学-分子动力学模拟软件包，是目前材料研究中常见的商用软件之一。更多详情见其官网说明：<https://cms.mpi.univie.ac.at/vasp/vasp/vasp.html>

³ 该工作流程部分参考 Jain A, Hautier G, Moore C J, et al. A High-Throughput Infrastructure for Density Functional Theory Calculation[J]. Computational Materials Science, 2011(50): 2295 - 2310，链接：<https://www.sciencedirect.com/science/article/abs/pii/S0927025611001133>

⁴ 价格参考自英特尔官网上各款处理器的价格区间：<https://ark.intel.com/content/www/us/en/ark/compare.html?productIds=192482,192446,192437,199346>

⁵ 更多戴尔易安信 PowerEdge 服务器详情，请参见戴尔官网介绍：<https://www.delltechnologies.com/zh-cn/servers/index.htm#section=rack-servers>

⁶ 该架构图未标出平台中相关的管理节点、服务节点和应用节点

⁷ 更多戴尔易安信 PowerVault ME4 系列存储详情，请参见戴尔官网介绍：<https://www.delltechnologies.com/zh-cn/storage/powervaultme4.htm#scroll=off>

^{8,9} 更多戴尔易安信 Isilon 系列可扩展 NAS 存储详情，请参见戴尔官网介绍：<https://www.delltechnologies.com/zh-cn/storage/powerscale.htm#scroll=off>

¹⁰ 更多戴尔易安信 VxRail 超融合一体机详情，请参见戴尔官网介绍：<https://www.delltechnologies.com/zh-cn/converged-infrastructure/vxrail/index.htm#scroll=off>

^{11,12} 以上数据援引自中科院物理所内部数据，如欲了解更多信息，请联系中科院物理所，<http://www.iop.cas.cn/>

有关中科院物理所材料数据库的更多信息，请访问：<https://atomly.net>

英特尔并不控制或审计第三方数据。请您审查该内容，咨询其他来源，并确认提及数据是否准确。

英特尔技术特性和优势取决于系统配置，并可能需要支持的硬件、软件或服务得以激活。产品性能会基于系统配置有所变化。没有任何产品或组件是绝对安全的。更多信息请从原始设备制造商或零售商处获得，或请见 intel.com。

没有任何产品或组件是绝对安全的。

描述的成本降低情景均旨在在特定情况和配置中举例说明特定英特尔产品如何影响未来成本并提供成本节约。情况均不同。英特尔不保证任何成本或成本降低。

英特尔、英特尔标识以及其他英特尔商标是英特尔公司或其子公司在美国和/或其他国家的商标。

©英特尔公司版权所有