

选择合适的工作站

如何选择合适的计算设备来处理图形密集型或计算密集型工作负载？本白皮书将探讨这背后的要素和原理。

Dell, Inc.
性能工程部
Alex Shows

选择合适的工作站

在配置和购买合适的工作站时，最重要的因素就是了解将如何使用它。预期用途决定了哪些组件对性能至关重要，哪些组件是可选甚至不必要的。此外，您越了解将如何使用工作站，您所能获得的性价比也就越高。先确定各种使用模式，再权衡这些任务的重要性和频繁程度，这样您就可以更有效地判断出适合相应工作的工作站。选择合适的系统配置，然后激活 Dell Performance Optimizer 2.0 新版软件，就可以对硬件进行调优，从而确保工作站发挥出最佳性能。

是计算型还是交互型

第一步是了解要在工作站上完成的工作类型，从而将主要任务划分为两种类别：计算型和交互型。

计算型

计算型任务通常是用户为分析某种数据模型而设置并随后运行的庞大而复杂的作业。在大多数情况下，这类任务所需的用户交互少之又少，它们所表现出来的特征是，系统的所有可用资源都会被它们大量占用。例如，视频帧渲染、买权/卖权预测、集成式有限元分析、蛋白质折叠、运动模拟以及新型赛车扰流板设计的下压力计算全都属于计算型工作负载。下面的图1是SolidWorks Simulation中的热流模拟示例。在该例中，计算型工作负载的运行结果得出后会转换成可交互的形式，从而使用户可以及时检验这些结果应用于模型的过程。

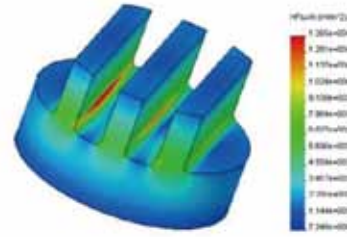


图1- 热流模拟

交互型

交互型工作负载需要用户进行大量的交互，这类工作负载的特点是，会零星地出现系统组件被大量占用的峰值时刻，在这些峰值时刻之间则是空闲时段，此时用户正思考下一项交互。交互型工作负载的示例有：查看和旋转发动机模型、对贯穿多层建筑的暖通空调管道走线进行批注，以及在三维建模程序中为复杂、完整工装的模型添加动画效果。下面的图2是一个在CATIA中操作复杂机械模型的示例。在该例中，整个总成的所有零件都可以在一个交互窗口中进行编辑和批注。



图2- 交互式建模

组件选择

通过将使用模式划分为计算型和交互型这两大类，有助于确定最优的组件配置（例如最佳的处理器、内存容量和插装的内存通道数量），以及这些组件的各种特性（例如可以达到的峰值CPU频率）的重要性。

对于繁重的计算型工作负载，将资金花在具体执行这种工作的相关组件（例如CPU和GPU）上最符合经济效益。此外，多路平台可以减少完成一项任务所用的时间（前提是处理这项工作的软件在性能上能够随处理器数目的增加而扩展），因而可以显著改进性能。如果应用程序因体系结构或许可方面的限制而不能跨可用的处理器进行扩展，那么再增加一路只会徒增成本和复杂性，因而可能是不合理的做法。



图3- 权衡核数与频率

与是否有必要再增加一路CPU的问题类似，有些计算型工作负载可能使用图形处理单元(GPU)作为计算资源，以此方式来扩展性能。将GPU想象成一辆短程竞速赛车可能有助于理解。在提供了一组数据（燃油）和一条直线跑道（可预测的重复指令）的情况下，GPU可以达到惊人的直线速度。另一方面，CPU则像是一辆拉力赛车。这辆拉力赛车内部的导航仪就像是CPU的分支预测算法，提示驾驶员接下来有哪些转弯以及如何以最好的方式顺利通过它们，而驾驶员则擅于快速应对高度复杂的车道上的各种路况。与之类似的是，很多计算密集型应用程序则擅于想方设法来通过使用GPU提高它们的性能。因此，务必要确定您的应用程序能否利用GPU，以及可能需要何种GPU。

CPU

在选择CPU时，首先要考虑有多少时间将用于处理计算型工作负载；处理这类工作负载时，所有可用的核心都将以很高的利用率长时间运行。在这种使用类型上花的时间越长，花在最大限度增加核数上的工作站预算比例就应该越高。首先要最大限度增加单路内的核数，同时还要考虑其他组件的预算要求。然后，如果需要更高的计算性能，则需要考虑改用配备双路CPU的平台，以便进一步提高计算性能。

一定不要首先就考虑使用双路CPU平台来最大限度提高计算性能。虽然这类平台能够提供最佳性能，但需要为此稍稍付出代价，这是多路体系结构的性质使然。这种代价会稍微降低显卡可以达到的帧率，从而影响交互使用模式。下面的图4说明了至强CPU体系结构对显卡性能的影响。至强E3级CPU通常采用最新的微体系结构并具备更高的频率。至强E5级CPU通常比同期的E3级落后一个微体系结构版本或一个微缩制程。最后，双路至强E5级在处理单线程工作负载（例如为GPU馈送输入的交互式应用程序）时会产生轻微的性能损失。有关显卡性能的更多信息，请参见下文。

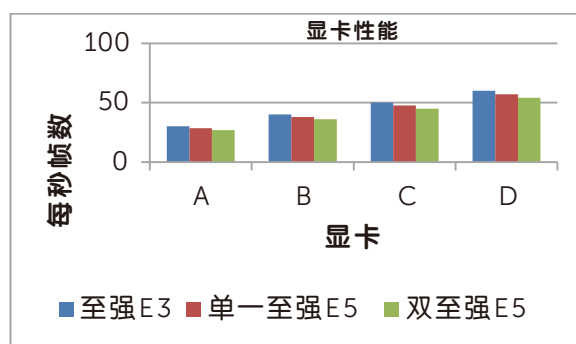


图4- CPU对显卡性能的影响

对于计算型工作负载，最大限度增加核数是最佳的做法；而对于交互型使用模式，提供最佳性能的方式则是使用最高的CPU效率。这是因为交互度（以每秒帧数加以衡量）常常受限于单个核心向GPU馈送指令和数据的效率。目前的大多数现代图形编程接口都只能使用单个线程向GPU馈送数据和指令，即使GPU驱动程序采用多线程处理也是如此。这就导致在核数超过四个时，增加核数所能带来的性能提升微乎其微。在交互型使用模式上花的时间越长，花在提高最高CPU频率上的工作站预算比例就应该越高。

Precision工作站中采用的大多数英特尔CPU都支持一种称作“睿频加速”的功能。在睿频加速模式下，CPU会根据其各个核心上分配的工作负载来调整其频率。当只有较少的核心处于忙碌状态时，CPU就会以较高的频率运行。当只有一个核心处于活动状态时，就可以达到最高的睿频加速频率。当很多甚至是所有核心都处于活动状态时，则采用最低的睿频加速频率。得益于这种动态的时钟频率控制机制，交互型工作负载能够以峰值睿频加速频率运行，而计算型工作负载则仍然以高于CPU标称频率的频率运行。这一点非常重要，因为比较两个CPU的标称频率（或它们的“额定频率”，通常会在型号名称旁标出）并不总能代表它们在绝大多数时间的运行频率。

为了更精确地比较CPU，应比较低频模式(LFM)、高频模式(HFM)、最低睿频加速频率（所有核心都有负荷时）以及最高睿频加速频率（一个核心有负荷时）。如果您希望在空闲时更节能，那么比较LFM就非常重要。如果CPU不执行任何处理工作，那么尽可能减少CPU的功耗有多么重要？

如果CPU不支持睿频加速，那么比较HFM就非常重要。如果CPU大部分时间都在运行计算型工作负载，那么比较最低睿频加速频率也很重要。最后，如果CPU大部分时间都在运行交互型或其他单线程工作负载，那么比较最高睿频加速频率就很重要。

在权衡是最大限度增加CPU核数，还是尽可能提高CPU频率，亦或在一定程度上兼顾这两者时，用户始终都应追求最新的CPU微体系结构和最新一代的CPU。较新一代的CPU通常采用制程微缩（较小的晶体管）或新型体系结构。较新的体系结构在相同频率下往往能带来较高的性能，而且这种改进不仅仅体现在CPU性能上。由于很多应用程序都会有一些时间浪费在等待单个核心向GPU馈送指令和数据上，因此随着CPU的整型运算性能提升，显卡性能也会相应提升。这意味着，采用较新一代体系结构的相同频率CPU在使用相同显卡的情况下可以达到较高的每秒帧数！

显卡

一般而言，对于显卡，价格越高，速度也就越快。显卡速度通常与以“每秒帧数”衡量的实时渲染性能相关。应用程序中的每秒帧数越高，您与数据模型的交互就越顺畅，您的工作效率也就越高。计算能力暂且不谈，要想为工作站找到合适的显卡解决方案，就需要考虑显卡在处理您最注重的应用程序时需要达到的每秒帧数。

对于显卡性能，一条值得借鉴的经验法则是，寻找在使用与您日常使用情形最接近的数据模型和渲染模式时，能够以每秒超过30帧的性能处理最重要应用程序的显卡。虽然视觉暂留现象表明，要保持流畅的动画视觉效果，至少需要达到每秒25帧，

每秒帧率越高，效果也就越好。如果特定显卡在采用特定的渲染方法以及特定的模型大小和类型时能够达到每秒100帧以上，那么就有理由认定，可以增加该模型的复杂度和/或大小，同时仍能够与该模型交互而不会出现卡顿现象。

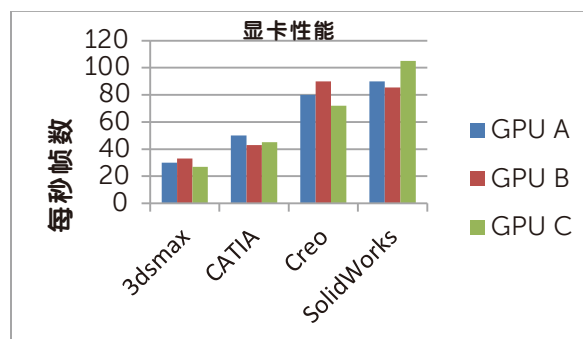


图5- 显卡性能取决于很多因素

上面的图5说明，显卡性能可能会因很多因素而异。要想选出最合适的显卡，最佳的做法就是评估显卡在处理工作站上使用的一款或多款特定应用程序时所达到的性能。虽然在升级到较高等级的显卡（核心时钟频率更高、显存更快、GPU核心更多，等等）时显卡性能通常会有大幅提升，但用户不能仅着眼于多款应用程序的总性能，单凭这一个因素就作出决定。这是因为同一等级的两款显卡在处理同一应用程序时所达到的性能水平可能大有不同。事实上，同一等级的显卡在处理同一应用程序时，仅仅改变数据集的复杂性或渲染模式，性能可能就会大有不同。推荐的方案是先为工作站选定GPU等级，然后测评该等级内的每一款显卡，以确定哪款显卡在您的特定使用情形下可以发挥出最佳性能。不过，对于大多数人来说，这并不现实，因此务必要使用某种其他标准测评方法作为替代。

SPECviewperf（在SPEC.org上提供）是一种用来比较不同工作站显卡的绝佳基准测试工具，因为它可以测评多种不同工作负载的每秒帧数，而测评时使用的渲染方法则可以反映很多常用工作站应用程序的渲染方法。利用这款基准测试工具，任何人都可以查看多种不同渲染方法的每秒帧数详细测评结果，并根据所公布的结果来比较显卡性能。这款基准测试结果还提供了可代表这些方法的图形质量的屏幕截图。如果是Creo用户，则可以使用这些数据来比较不同显卡之间的性能，不仅仅可以比较在处理Creo时的性能，还可以具体比较在处理最能代表他们特定Creo使用情形的数据模型和渲染模式时的性能。

在考虑哪款显卡最适合您的工作站时，需要衡量该工作站在典型的一天里有多少时间用来处理高度交互性的工作或用来处理会用到GPU的计算型工作。在后一种使用类型上花的时间越长，花在显卡上的工作站预算比例就应该越高。相反，在后一种使用类型上花的时间越短，花在CPU、内存、存储等其他组件上的工作站预算比例就应该越高。

内存

据说，随机存取存储器(RAM)永不嫌多。虽然对于运行大量采用多线程处理的应用程序的现代多核系统来说，这句格言可能也适用，但在考虑要为工作站配备哪种内存时权衡其他因素仍是十分重要的。对于计算型工作负载，几乎时时刻刻都需要最大限度增加可供处理核心使用的内存带宽。因此，如果可以选择是插装八个各有8GB容量的DIMM，还是插装四个各有16GB容量的DIMM，请选择插装更多DIMM插槽的方案。增加可用内存带宽会降低内存

带宽成为计算型工作负载瓶颈的可能性，从而将计算负担转移回CPU核心、频率和高速缓存。

选择合适的频率也很重要，具体频率因工作负载而异。在需要最高内存带宽的应用程序中，用频率最高的内存插装用频率最高的内存插装十分重要。

不过，有些应用程序不论可用带宽如何，都需要尽可能降低延迟，在这种情况下，则需要用频率较低的内存插装所有可用的DIMM插槽。例如，如果需要随机访问足够小、适合用作CPU高速缓存的内存，就属于这种情况，但CPU无法预测接下来要访问的具体内存位置。

虽然内存带宽依然重要，但较慢内存速度所具备的较短延迟可以让这些随机读取和写入操作受益。举例来说，在高速买卖交易等金融市场，监控股票和大宗商品等投资对象的现价的应用程序就需要最低延迟，以便它们可以尽可能快地对市场变化作出反应。由于在零点几秒之间就可能赢得和损失巨额财富，因此，即使总体内存带宽较小，频率较低的内存所具备的较短延迟也可以带来优势。

下面的图6以SPECwpc计算型工作负载为例说明了这一概念。在所有其他方面都相同的情况下，配备双至强E5 CPU的Precision T7610平台在DIMM数目增加时得分显著提升。这意味着在真实环境中完成渲染、有限元法(FEM)分析、计算流体动力学等诸如此类的工作所用的时间将减少。应注意的是，这些益处仅限计算型工作负载。不过，插装更多DIMM插槽给交互型工作负载带来的益处则很

难衡量，因为大多数图形工作负载都可以使用显存加以处理且大多数显卡上都配有专门的显存。

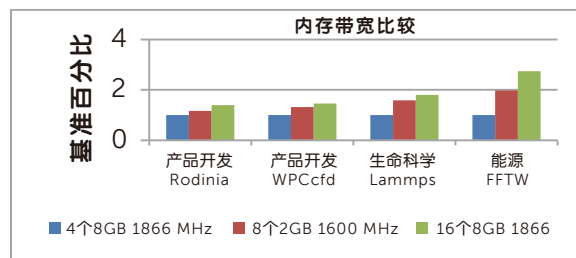


图6-通过增加插装的DIMM插槽数实现的SPECwpc改进

最后，当各项计算中使用的数据的完整性对于最终结果至关重要时，应使用差错校验(ECC)内存。例如，对一个大型数据集进行迭代时，计算输出不断地作为输入提供给另一计算序列，这种情况下早期计算中遗漏的一项错误可能会对最终结果产生显著影响。

Dell Precision工作站的一项独有特性是，纳入了戴尔独创且已获专利的可靠内存技术(RMT)。如果出错，RMT会根据错误纠正码(ECC)数据来找出DIMM中出错的具体位置。RMT可以发现位于特定内存位置的错误，然后自动在内存中掩蔽该位置，以确保后续的读写操作使用正常的位置。这种机制的最终效果是，当ECC错误大规模出现时，不必更换整个DIMM，RMT只需掩蔽很小的相关区域即可让DIMM正常运行，从而提高了系统的可用性和可靠性；而且，由于用户可以继续使用原本在其他系统上需要立即更换的DIMM进行工作，内存的使用寿命也得以延长。

存储

在存储性能方面有各种各样需要考虑的事项，具体完全取决于使用模式。例如，数据是在网络上还是存储在本地？如果是在网络上，多久需要向网络资源提交一次更新？如果是存储在本地，需要在本地配备多少容量？是否需要为本地存储建立冗余机制？所有这些因素对于确定适合工作站的存储组件而言都十分重要。考虑到这种复杂性，虽然本文对这一主题仅稍作简述，但实际上在这方面需要格外注意。

为简单起见，我们假定数据存储在工作站本地，因而不需要考虑网络带宽、更新频率以及签入/签出过程。一名工作站用户将面临以下三种常见的本地存储使用情形，三种情形或兼而有之：

- “**办公效率**”——读写小型文件，偶尔会传输大型文件
- “**交互型工作站**”——打开和保存各种大小的文件
- “**计算型工作站**”——对非常庞大的数据集进行迭代，常常会产生很大的临时文件

针对“办公效率”这种使用情形进行优化通常很简单：只需在预算范围内权衡预期的容量需求与性能最高的驱动器类别即可。虽然旋转型驱动器以往在这一细分市场占据主导地位，但近年来随着MLC（多层单元）内存和控制器的成本不断降低，更为实惠、令更多用户都能承受的固态硬盘(SSD)和混合硬盘应运而生。一般而言，针对这种使用情形，混合硬盘可以带来最佳的性价比，而固态硬盘则可以立竿见影地带来最优性能。混合硬盘的运行机制是，将最常用的数据存储于缓存中，因为从缓存中访问比从驱动器内的旋转型介质

中访问更快。只要处于使用中的文件相对较小，数据就会保留在闪存中，从而可以加快性能。

下面的图7说明了在各种使用模式和驱动器类型之间可以进行的典型扩展。SATA与串行连接的SCSI (SAS)驱动器在体系结构上非常相似，因此它们的扩展是在工作负载间均匀进行，并且主要取决于磁盘接口类型、旋转速度以及板载内存。混合式磁盘的性能可能随着工作负载的不同而有显著差别。工作负载的确定性和重复性越强，混合磁盘的性能也就越高。不过，混合磁盘受限于其闪存缓存的大小，因此对于以不确定方式对大型数据集进行迭代的计算型工作负载，混合磁盘所带来的益处就不如“办公效率”使用情形。SATA和PCIe固态硬盘相比旋转型驱动器具备显著优势，在随机读取方面这种优势最为明显。SATA固态硬盘受限于其接口类型，因此如果要使互动型和计算型工作负载达到最大吞吐量，采用PCIe固态硬盘可以带来显著改进。

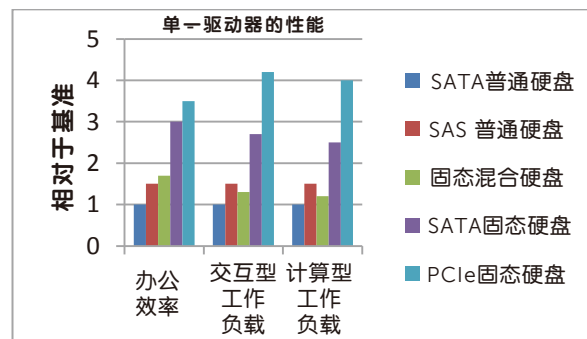


图7-单一驱动器性能比较

“交互型工作站”使用模式需要更高的性能，也正是在这种使用模式下，固态硬盘、SAS驱动器和RAID阵列开始发挥更重要的作用。如果单个固态硬盘可以同时满足您的办公效率和交互型工作站使用模式的容量需求，那么此方案将是除多驱动器RAID 0以外性能最佳的方案。

利用RAID阵列，可以创建跨一个或多个物理（或逻辑）驱动器的大型虚拟驱动器。根据RAID类型，还可以实现冗余（同时有多份数据副本）等新功能并提高性能。如果冗余比性能重要或与其同等重要，那么采用RAID阵列（例如RAID1、10或5）将是更好的选择。这样的话，需要决定的就是在构建阵列时有多少驱动器可用（匹配）。改用RAID可能会大幅提高存储成本，导致无力承受在阵列中纳入高性能驱动器的成本。要想在保持高性能的同时降低这种成本，一种方式就是使用包含操作系统和应用程序的固态硬盘引导驱动器，同时使用成本较低的旋转型磁盘驱动器构建RAID阵列来存储较大的数据集。

对于计算型工作站，由于在这种工作站上会使用极大的数据集，因此唯一适合这类使用情形的方案可能就是由大型驱动器组成的RAID阵列。通过将容量较小的驱动器组合成单个大型卷，应用程序可以像使用单个大型驱动器那样使用所有这些容量。RAID 0中的多个驱动器可以最大限度提高性能和容量，但它不会提供任何冗余。RAID 1中的多个驱动器可以提供冗余，但无法最大限度提高性能或容量。RAID 10可以提高性能、容量并增加冗余，但如果从需要的驱动器数量方面考虑，其成本最为高昂。

在RAID 0与RAID 10之间是RAID 5，它能够以比RAID 10更少的驱动器提供更高的性能、容量并增加冗余，但由于需要计算分布在阵列范围内的奇偶校验数据，因此需要耗费更多的开销来管理阵列。在考虑是否要向集成式存储控制器添加第四个驱动器以及在创建RAID 10时，可以考虑升级到配备板载内存的独立RAID控制器的方案并改用RAID 5。您可能会

看到更高的容量，而且添加独立RAID控制器可能意味着在办公效率和交互型工作站使用类型中能够实现更高的性能，更何况还会给计算型工作站使用类型带来益处。

下面的图8说明了使用不同的RAID类型跨多种使用模式进行的性能扩展。这里的扩展采用一张完整的硬件RAID控制器卡，这样奇偶校验计算和整个RAID协议堆栈的负荷就会从CPU转移出去。系统配置方面的所有其他可变因素都保持恒定。每个条也是多个不同衡量标准（包括读取、写入、随机、顺序以及这些标准的变化形式）的平均值。如您所见，包含两个驱动器的RAID 0性能很好，但无法提供冗余。RAID 1可以提供冗余，而且因包含两个驱动器而确实可以提高读取性能，但每当必须向磁盘提交写入数据时，写入操作都会出现轻微的性能损失。RAID 5凭借三个驱动器可以带来显著的优势，而且硬件RAID控制器使CPU不必承受奇偶校验计算负担，从而减轻了对这些操作的写入性能的影响。如果采用的是软件RAID控制器，一定要避免使用RAID 5，否则CPU将承受奇偶校验计算负担，并且在很多情况下性能都将大大低于其他RAID类型。RAID 10虽然需要使用四个驱动器加以构建，但它非常好地兼顾了性能与冗余。

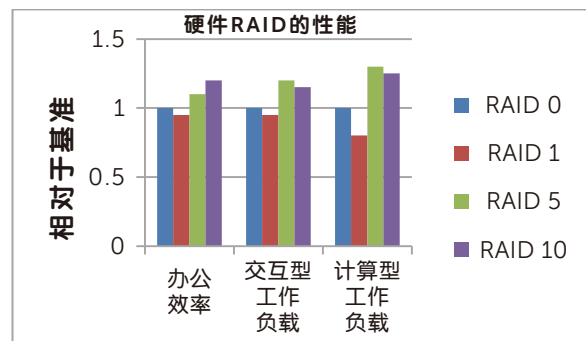


图8- 硬件RAID的性能

应用程序认证

专业工作站相比传统PC所具备的其中一项关键差别性优势是，它们获得了各种平台认证，从而可以运行特定的专业工作站应用程序。考虑到软件环境的复杂性，不难想象，在操作系统版本、应用程序版本以及硬件、固件和驱动程序版本方面可能存在的变体数量惊人。所有这些可变因素都可能会影响应用程序的稳定性和性能。工作站认证机制可以仔细记录系统的哪些配置已确定与相关的特定应用程序版本兼容，从而化解这种复杂性。

Dell.com/Workstations上提供了最新的认证列表。这可以降低用户在考虑是购置新工作站还是升级现有工作站时所面临的风险，因为他们在购买前可以确信，特定工作站已通过他们所需的应用程序的软件供应商认证。

总结

要找到合适的工作站配置，请先确定第一位的使用模式，以及任何第二位或第三位的使用模式。

对于交互型使用模式，应侧重于最大限度提高CPU频率，然后是显卡等级。对于各个CPU型号，应选择采用最新体系结构的型号，并且应主要考察峰值睿频加速频率。面对可供选用的各种CPU型号，应确定频率与价格之比最优者。然后考察显卡，并使用SPECviewperf等行业标准基准测试工具比较每秒帧数，重点测试对您的使用情形最为重要的应用程序和/或渲染模式。

判断可供选用的GPU型号中哪种型号所达到的每秒帧数与其价格之比最优。然后考察内存，确定在提供所需容量的同时能够带来最高内存带宽者。最后考察存储；在这方面，单个固态硬盘或许能满足交互型使用模式的所有需求，除非出于容量或冗余考虑而需要使用RAID阵列，或者因支出限制而只能选择单个旋转型磁盘驱动器。

对于计算型使用模式，应侧重于最大限度增加核数，然后是CPU频率。对于各个CPU型号，应选择采用最新体系结构的型号，并且应主要考察最低睿频加速频率（此频率反映CPU在重负载下可以通过睿频加速技术达到的最低频率）。优先选择核数与价格之比最优者；如果工作站的整个使用寿命中将有超过一半的时间处理计算型工作，可以考虑升级到双路工作站。如果应用程序支持GPU计算，可以考虑将GPU升级到配备更多计算核心的型号，因为采用GPU升级做法的性价比往往要比CPU高（这里也是根据核数增加百分比来衡量的）。先通过插装尽可能多的插槽来升级内存；除了为数不多的对延迟高度敏感的应用程序之外，最佳的做法往往是升级到最快的内存速度以尽可能提高计算吞吐量。最后要考虑的是存储需求，主要从应用程序所需的容量和带宽着眼，计算使用模式的这两项往往远高于其他使用模式。

要确定特定使用模式的最优性能，可以找出对这种使用模式最重要的因素，也就是对性能影响最大的那些因素。如果一款工作站满足上述选择要求，通过了相关认证，并且装有所需使用模式会用到的关键应用程序，那么它就可以确保用户为它支付的任何价格都能带来最佳的体验。