

全闪存阵列技术手册

全闪存阵列技术手册

固态存储（SSD）技术已逐步渗入到服务器、混合存储阵列以及缓存设备的应用中，不过，新兴的全闪存阵列正在开始进入高性能存储系统市场。

全闪存存储系统是完全由固态存储介质（通常是 NAND 闪存）构成的独立的存储阵列或设备。这些系统是用于增强可能包含磁盘阵列的环境的性能，或者用于取代所有传统的硬盘存储阵列。并且由于其便捷性、在总体拥有成本方面的竞争力以及高性能等优势，使得其越来越受到企业和厂商的关注。

但是，固态存储阵列并不能实现传统阵列那样的高扩展性，并且相对昂贵的 SSD 使得全闪存阵列只适用于高 IOPS 应用环境，而非通用存储。但是，由于其在高性能方面的优势，全闪存阵列一定会在绝大多数企业级环境中找到一席之地。

本技术手册分析全闪存阵列技术。

- ▣ 步入全闪存 SSD 存储系统时代
- ▣ TechTarget 新年预测：六大热门技术之全闪存阵列
- ▣ 全闪存存储系统的类型和功能
- ▣ 全闪存阵列解析：SSD 的类别和缺陷
- ▣ 全闪存阵列解析：SSD 设计及全闪存应用案例
- ▣ SNW：SSD 用户探讨混合阵列及全闪存阵列
- ▣ 只要使用正确 混合式和全 SSD 阵列都可以节省费用

步入全闪存 SSD 存储系统时代

已经有人开始揣测固态存储将在企业级存储系统中占据主导地位。这种推断背后的依据是闪存 SSD 相较于磁盘驱动器具备更高的 IOPS、更低的能耗、降低制冷需求和更低的故障率，因此有人才预估闪存 SSD 将最终在存储系统中取代磁盘驱动器。然而分析师一直以来（今后也将会）都坚持认为闪存 SSD 单位容量代价高昂，并且从以往来看，事实也确实如此。

成本因素使得存储系统供应商开始着手在提升闪存 SSD 优势的同时降低其价格。例如，将闪存 SSD 作为缓存或第 0 层存储，并配合自动化分层软件。这种解决方案通常被称为混合存储。

缓存方式加快了读取操作，同时在某些情况下可以将重要数据保留在闪存 SSD 的缓存中。闪存缓存的方式主要基于智能的读取频率分析。所有的写操作都直接写入磁盘驱动器中。基于读取的频率，数据会被放置到 SSD 缓存上。当读取频率下降，数据就从缓存中移除。用这种智能闪存缓存面临的问题是在于策略上的限制和缓存总量。同样，过少的缓存会导致闪存丢失以及重定向到磁盘驱动器的次数，而这和闪存缓存的初衷相悖。

第 0 层存储使用闪存 SSD 为需要高性能的应用程序提供主存储，尤其是那些需要高 IOPS 的应用。要使得这种方式更为高效，需要基于策略的自动化存储分层技术（例如不需要管理员参与将数据从一个存储层迁移到另一个上）。自动化分层基于数据使用价值。随着数据的老化，其使用价值下降，很少被再次使用。使用第 0 层自动化存储分层技术的问题在于其自动化存储分层软件的成本和系统复杂度。软件成本在很大程度上可以从降低闪存 SSD 的部署数量上得以弥补。自动化分层对存储管理员是很有吸引力的；不过部署过程要求相应的技能和对细节的关注，例如确保快照和数据保护策略以及数据的匹配度，而无需担心数据迁移到那一层级。

在存储系统内不建议过多的部署闪存 SSD 的一项潜在原因是考虑到存储控制器和背板的性能局限。当系统内部署大量闪存 SSD 以后，这些就变成了性能瓶颈。因为在设计之初就没有考虑到闪存 SSD 性能。因此，混合优化闪存 SSD 和磁盘存储系统变成了新的主流。

不过如果我们假设闪存 SSD 可用容量成本和 10000 转 SAS 磁盘一样又会如何？而且如果存储系统能够完全利用大量的 SSD 所产生的高性能，结果又会怎样？这些问题的答案即将从一些新兴的存储系统供应商那里得以解答，譬如 Astute 网络、Nimbus 数据、PureStorage、Skyera、SolidFire、Tegile、Texas 内存系统（现已被 IBM 收购）、Violin、Whiptail、XtremIO（现已被 EMC 收购）以及其他一些公司。

这些供应商能够提供（或者已经在未来的产品中规划，例如 EMC-XtremIO）全闪存 SSD 存储系统。他们可以确保其存储控制器体系架构完全使用闪存 SSD 的性能。有一些使用纵向扩展的设计结构，例如 Nimbus 和 PureStorage。另一些则会使用横向扩展的架构，其中包括 Skyera、SolidFire 和 EMC 的 XtremIO。此外，还有一些公司更加关注在一些细分市场上，其中有关虚拟机存储优化的 Astute；确保云服务供应商 SLA 性能管理的 SolidFire；针对 HPC 市场提供极高性能应用的 TMS 和 Violin；以及关注中低端市场的 Whiptail。

通常部署全闪存 SSD 系统的前期成本从每 GB 3 美金到每 GB 15 美金不等。这一成本是在存储系统执行内置的重复数据删除技术和压缩技术之前得出的。全闪存 SSD 系统的性能可以承受得起在应用的同时进行重复数据删除和压缩。（在混合存储系统中也同样如此，但对磁盘存储系统就不一样了）因此实际每 GB 的价格范围大约在 1 美金到 5 美金之间。

相比之下，闪存 SSD 和磁盘的混合系统在重复数据删除和压缩之前，其前期成本从每 GB 1.6 美金到 24 美金不等，重复删除和压缩后的则大约在 0.54~8 之间。而传统的磁盘存储系统前期导入价格在每 GB 0.5~15 美金，其中取决于 SATA、近线 SAS、SAS 和 FC 磁盘的混合情况。

	Avg upfront per GB HW Cost	1yr avg per GB Power & Cooling	Avg upfront per usable GB Costs*	~ 3 yr per usable GB TCO
100% HDD Storage System	\$0.50	\$0.50	\$0.63	\$2.13
Hybrid Storage System	\$1.60	\$0.38	\$0.59	\$1.72
100% Flash SSD Storage Systems	\$10.00	\$0.17	\$3.64	\$4.13
100% Flash SSD Storage Systems	\$5.00	\$0.17	\$1.82	\$2.31
100% Flash SSD Storage Systems	\$4.00	\$0.17	\$1.45	\$1.95
100% Flash SSD Storage Systems	\$3.00	\$0.17	\$1.09	\$1.59

* HDD cost/usable GB goes up because of formatting and RAID overhead.
 SSD based dedupe/compression increases usable GB ~ 2.75 x.
 Hybrid systems a little less to account for the HDD overhead.
 SSDs are typically overprovisioned from 20 to 50% to account for load balancing & garbage collection

TechTarget | 中国

如果考虑到电力和制冷因素，全闪存 SSD 系统成本则更具竞争力。闪存 SSD 存储系统的供电和制冷消耗相对于磁盘系统降低了大约 50%~80%。跟踪显示，每年磁盘存储系统的供电和制冷成本和前期硬件的导入成本相当。比较全闪存存储系统和混合系统的供电和制冷取决于 SSD 和磁盘的混合情况，以及磁盘是否支持 MAID 技术。为了方便比较，我们大致上假设相比于 100% 的磁盘系统，能耗和制冷费用只需 25%。同时假设普通 SSD 系统的供电和制冷是磁盘系统的 67% 或三分之二，那么对于混合系统就差不多是 50%。（75% x 67% = 50%）而前期每 GB SSD 的成本大约为 5 美金，整体总体拥有成本的比较相当有意思。

随着硬件成本下降到 5 美金以下，SSD 系统 3 年内的平均每 GB 总体拥有成本将变得很有竞争力。但是在这种比较中有太多的假设，不过这也显示了 SSD 和 HDD 的价格差距变得如此接近，超过绝大多数人的预期。

当我们回到最初的那个问题：随着全闪存 SSD 存储系统各项障碍的消除，这种转变也已经开始了么？基于最近的厂商销售报告，答案是肯定的。一家品牌时间较长的全 SSD 存储供应商 Nimbus 数据的报告表示年销售增长达到了 500%。另外一些，例如 Astute、Violin 和 Whiptail 的销售业绩增长也超过了传统的存储系统。当然这些报告存在一些值得商榷的地方。所有这些厂商都是私有制的，不需提供相应细节。同时这种增长的基数十分渺小。但这些数据同时也指出了未来的方向。随着时间的推移，我们会清晰地看到这是一种行业内的转变，还是一种短时间的异常现象。

最后，全闪存存储系统，由于其便捷性、在总体拥有成本方面的竞争力以及更高的性能，为这种系统带来了新的存储商机。还有一点值得关注的是，许多传统存储供应商所提供的全闪存存储系统价格远高于新兴企业。因此不要迷信传统的存储供应商。

(来源：TechTarget 中国 作者：Marc Staimer 译者：张瀚文)

TechTarget 新年预测：六大热门技术之全闪存阵列

TechTarget 2013 年热门技术预测新鲜出炉了。重点关注的技术并不是实验室中的最新技术，而是已经获得一定应用的技术，固态存储、云存储、虚拟化和数据保护都位列榜单之上。

本文我们一起谈谈全闪存阵列。

价格一直是全闪存阵列的绊脚石，但是一些初创厂商将 SSD 价格拉低，使得全闪存阵列的应用成为现实，一些巨头对闪存厂商的收购使得闪存更深入的进入企业级市场。

简单而言，对速度的渴望使得固态存储非常有市场。顶级全 SSD 阵列可以提供 500,000 到 1,000,000 IOPS，即使是“二流”全闪存阵列也能够提供 100,000 到 200,000 IOPS。Taneja 集团高级分析师 Jeff Byrne 说，“当你考虑使用全 SSD 阵列时，你希望以最小的投资获得最高的 IOPS 或存储性能。”

按照每 GB 的价格计算，全 SSD 阵列确实价格不菲，但如果从每 IOPS 的价格计算，确实还是物有所值。所以全 SSD 阵列的应用最佳场合便是需要高性能的场景。

Byrne 表示，“数据分析、数字成像、VDI、数据库应用、金融交易系统以及游戏网站等都适合使用 SSD”，这些应用需要非常高的传输速度和非常高的随机 I/O，使用昂贵的 SSD 也算是物尽其用。

全闪存平台多来自于一些初创厂商，如 Kaminario、Nimbus Data、Pure Storage、SolidFire、Skyera、Tegile Systems、Violin Memory 和 Whiptail Technologies。一些大厂商也紧随其后进入了这个市场，去年五月份 EMC 收购了 XtremIO，并计划 2013 年年中推出全闪存 Project X 系统。IBM 也在 2012 年 8 月份收购了 Texas Memory Systems，进入了全闪存市场。

前几年，顶级厂商的全闪存阵列价格约在 16 到 20 美元每 GB。中级阵列的价格也在每 GB 3 到 8 美元。而现在企业级闪存的价格则低于每 GB 2 美元。

全 SSD 存储厂商一直在致力于降低阵列的成本，同时保证整体阵列的设计更加高效。

降低成本的方法之一便是采用 MLC 而不是更加昂贵的 SLC。SLC 耐性和可靠性更强，可以支持 10 万次的擦写。MLC 大约支持 1 万次的擦写，但厂商通过使用软件和更好的写入技术提升了 MLC 的性能和可靠性。数据减除技术也可以通过有效的提升存储空间实现价格的降低。

随着闪存价格的降低，2013 年 MLC 的应用将更加广泛，我们将看到闪存走入传统的企业级应用环境，在某些情况下将取代机械硬盘。

(来源：TechTarget 中国 译者：吴迪)

全闪存存储系统的类型和功能

全闪存存储系统是完全由固态存储介质（通常是 NAND 闪存）而没有硬盘驱动器（HDD）构成的独立的存储阵列或设备。这些系统是用于增强可能包含磁盘阵列的环境的性能，或者用于取代所有传统的硬盘存储阵列。它们也是用于支持由于 NAND 闪存较佳的性能带来的特定需求和功能，以及克服对闪存存储独有的写/擦除周期的挑战。

在一个全闪存存储系统中，控制器被优化以获得 SSD 的高性能。传统的磁盘控制器通常每个可支持少于 200 IOPS 的存储媒介（磁盘驱动器）。高性能的 SAS HDD 几乎可以达到这个数字的 2 倍，但与一个普通的固态驱动器提供 8000 IOPS 相比仍然微不足道。考虑到磁盘驱动器的延时，HDD 阵列控制器可以处理周期剩余，大多数的处理周期用于处理快照、卷管理和复制之类的存储服务。

当 SSD 取代 HDD，即这个延时消失后，控制器成为瓶颈，而且其性能显著变差。全闪存存储系统推出保持数据流向较快的存储介质的控制器，无论这些存储介质是驱动器规格的 SSD 或是闪存电路卡模块形式的。它们通常具有允许更多数据流进出存储介质的分布式的控制器架构，经常会以独立的处理器处理存储服务的系统开销。有些也具有降低网络协议处理对性能的影响的板载（onboard）功能。

全闪存存储系统的控制器也处理耐久性问题。与磁盘驱动器不同，NAND 闪存无法在位级别进行覆盖写入。为了适应这种变化，闪存的一个块在新数据写入前必须整个擦除。这需要单独的内部管理程序，整合打算删除而从块里面保存出来的数据，在擦除前创建额外的数据拷贝步骤。所有这些被称作“垃圾收集”的额外的系统开销步骤消耗存储控制器的 CPU 周期，如果这些周期同时从处理数据 I/O 的处理器被偷走，会影响性能。

这些解决方案也无独特之处，全闪存存储系统拥有指定专门硬件处理这种闪存特有的系统开销的闪存控制器，因此不会降低性能。与基于软件的闪存控制器相比，它们运行这些 NAND 闪存操作（减少写入放大、提高损耗均衡）的工作更有效率。该功能也使低成本的多层单元（MLC）和企业级 MLC（eMLC）NAND 闪存芯片的使用成为可能，有助于降低全闪存存储系统每 GB 的实际成本。

现在我们讨论一下现有的全闪存存储系统的类型以及用户是如何使用的。

功能 1：增强性能

全闪存存储系统的第一个功能是作为对已有的基于磁盘的基础设施的高性能存储设备的增强。新技术最早的采用者通常是处于最前沿的那些人，那些似乎不惜代价需要更多性能的用户。现在，这些闪存设备成本已经下降，但通常只是用于增强性能。在这种环境里，全闪存系统能够为游戏或者高事务数据库等在线应用在负载高峰期的时候的运行提供了临时或兼任的块存储区域。EBay 就是使用全闪存设备增强这类环境的性能。

全闪存设备除了充当闪存层，也能够加速存储在硬盘阵列的应用，或者为虚拟服务器和 VDI 环境提供网络化的高性能高速缓存。Texas Memory Systems、Kaminario、Violin Memory、Whiptail 和 Astute Networks 公司已经推出这类产品，尽管其中很多公司也有其他闪存类别的基于闪存的产品。

功能 2：取代 HDD 阵列

全闪存存储系统的第二个功能是以全闪存阵列的形式取代已有的基于磁盘的存储阵列。这些解决方案的部署与第一组功能相同的基本技术，具有专为闪存存储设计的控制器和内部架构，但也加入了存储服务。作为传统的磁盘阵列的替代，全闪存存储阵列需要满足市场上对卷管理、快照和复制之类的功能的需求。有些也提供多协议功能的基于块和基于文件的存储。

全闪存存储阵列通常包含重复数据删除、自动精简配置和压缩之类的存储效率技术，因为这些程序在固态存储中工作良好。数据缩减也能够降低每 GB 的实际成本，这在过去一直是更广泛采用闪存存储设备的主要障碍。尽管为主要的存储阵列实施重复数据删除之类的技术可能会有性能影响，制造商们期待全闪存系统仍有性能剩余。在这个领域的制造商包括 Pure Storage、Nimbus、Skyera、Greenbytes、EMC（与 XtremeIO 合作）以及 Solidfire。如上所述，其中很多公司的解决方案可归类到全闪存设备，也可用于增强性能。

全闪存系统的主要优点

与其他闪存技术比较，全闪存系统一般具备几个主要优点。不需要使用服务器端的 SSD 或者 PCIe 卡，他们可以像目前的磁盘阵列一样在 SAN 配置环境中进行共享，这点对那些有许多潜在服务器主机需要加速的用户具有吸引力。共享对于提供高可用和支持 VMware Storage vMotion 的服务器虚拟化实施也具有吸引力。

与混合闪存系统（同时带有闪存和磁盘的设备）相比，全闪存存储系统提供更快、同时也更一致和更可预测的性能。与混合阵列不同，它们的控制器可以忽略决定哪些数据集应该放在闪存以及在两个存储区域之间实际往返移动数据所需要周期的高速缓存和分层运算法则的系统开销。全闪存系统的数据都置于固态存储内，在服务器虚拟化之类的高动态环境中，不像高速缓存或者分层系统那样容易受工作量变化的影响。不会有高速缓存或者层“丢失”的风险，这种破坏性的场景之下，所需数据没有如期在闪存上而从磁盘获取数据时应用被迫等待。

成本：TCO 问题

如上所述，闪存系统的每 GB 的成本问题一直抑制基于闪存的存储设备的采用。但当重复数据删除和其他数据缩减技术对传统的磁盘阵列不划算时，它们特别适合全闪存系统。首先，这些程序在全闪存系统上能运行良好，因为这些系统已有的高 IOPS 性能，使这些程序比在 HDD 上更有效率。同时，闪存的每 GB 基础成本比较高，所以，与 HDD 相比，数据缩减额外取得的每一个百分点，都有更好的边际回报。由于这个原因，计算全闪存系统的投资回报的时候，必须留意需要包括合理的数据缩减比率。

增值代理商的观点

闪存已经成为目前存储技术项目里的固定设备。不管是作为性能增强型设备进行实施或者取代传统的 HDD 系统，问题不再于是否使用基于闪存的系统，而是另外一个选择：当高性能磁盘驱动器和 SSD 之间每 GB 实际成本差别不断缩小时，这个选择越来越划算。且不说潜在的全闪存项目，增值代理商也应该熟悉全闪存存储系统，因为这项技术是闪存存储更大的话题的一部分。除了服务器端 SSD/PCIe 解决方案和混合闪存设备以外，全闪存系统也是客户将会问及的一个实施选择。

(来源：TechTarget 中国 作者：Eric Slack 译者：李志明)

全闪存阵列解析：SSD 的类别和缺陷

固态存储技术已逐步渗透入服务器、混合存储阵列以及缓存设备的应用中，不过新兴的全固态驱动器阵列正在开始进入高性能存储系统市场。

50 多年以来，磁盘驱动器一直作为主要数据的存储途径，但其地位在今天正在受到挑战。随着价格的下降以及稳定性的提升，固态存储开始逐步打入企业级市场；并且随着固态存储产品中增加的磁盘驱动器技术，出现了大量提供全闪存阵列产品的新兴企业。

虽然在传统的阵列中部署固态驱动器的方式已有数年之久，但这种简单地将磁盘驱动器替换成固态驱动器的方式却无法保证高性能。正如我们所能预见的，要实现固态驱动器和闪存的最佳效能，其处理方式和传统的旋转磁盘完全不同。

SSD 之间的差异

固态驱动器产品中包含了闪存和控制器电路，工作方式和通常的磁盘驱动器类似。现今该产品有两种使用类型，单层式存储（SLC）和多层式存储（MLC）。单层式存储驱动器在每个单元内存储一位字节，而多层存储驱动器能够在每个单元内存储多个字节（通常是两位）。就每 GB 而言，SLC 驱动器的价格更加昂贵，不过使用时间也更长并且通常较 MLC 产品而言性能更快。MLC 驱动器每 GB 相对便宜一些，不过在耐用性（在读写操作中出现不能修复的错误之前的使用时间）确实是一个问题。一些供应商同时也推出了第三类闪存，企业级多层式存储（eMLC），品质较 MLC 更好，提供更长的使用时间。

闪存可以包装成磁盘驱动器的样子，模拟标准的 SAS 或 SATA 协议。绝大多数供应商会选择使用这些接口用于 SSD 部署，不过虽然接口一样，SSD 和 HDD 的特性完全不同。

磁盘驱动器天生属于机械类，所需要的物理移动造成了 I/O 运作时的延迟。HDD 在处理高随机工作负载效率尤其低下，因为需要不断将读写磁头移动，不过在顺序读或写操作时，其表现不错。

而 SSD 没有要移动的部件，在顺序或随机 I/O 时都不会造成额外的负担，这使得 SSD 在某些场合下十分有效，比如虚拟桌面基础架构（VDI）。

固态硬盘较磁盘驱动器在读写方面都有数量级上的提升；一块通常的 SSD 驱动器即便在随机 I/O 的使用环境下也能够达到 100,000 的 IOPS。同时，较之于磁盘驱动器，SSD 的重量也更轻，并使用更少的能耗，对于许多企业存储阵列的总体拥有成本（TCO）也有不小的影响。

SSD 的缺陷

除了这些表面的优势，SSD 也绝非完美的驱动器，有其自身的缺陷。最为明显的是闪存的有限写操作周期，SSD 最终会损耗殆尽。通常，SLC 闪存的耐久度大约为 100,000 次写周期；MLC 闪存则少了一个数量级，每个数据单位只有大约 10,000 次。除了这些，数据的存储和读取也并不完全可靠。厂商方面应用了许多技术来扩展使用周期，包括损耗平衡、纠错编码、坏数据块重新映射以及过量配置。



IOPS/\$与 IOPS/GB

传统的每 GB 美元参数不再适合于全闪存阵列；在比较供应商的产品时，取而代之的单位美元的 IOPS 和单位 GB 的 IOPS 能够起到更好的衡量作用。在高 I/O 环境中寻求最佳性价比时，单位美元 IOPS 显示了设备的成本效益。而单位 GB 的 IOPS 则显示了一台阵列的相对 I/O 密度，即存储的每个 GB 能够提供的读写速度。

TechTarget | 中国

SSD 的响应时间同样有可能不一致。数据写入 SSD 时会送到整个驱动器上以确保每个独立单元均匀磨损。这一流程称为损耗平衡。随着新数据的写入，各单元块会被重新标记回收。随着一块 SSD 达到较高的利用率或写操作，该回收过程可能会因为无效数据区域的覆盖操作延缓响应时间。很显然，这种情形是低延迟环境中不可取的。

是否要效仿磁盘驱动器

将闪存包装成传统磁盘驱动器一样的形式有不少明显的好处。其可以简单地整合到盘阵中取代现有的磁盘驱动器，并且其支持现有的协议和形状。绝大多数 SSD 阵列供应商会这样选择，使用常见的 SSD 作为标准形状。不过，有两家厂商并没有这样做。

Nimbus Data System 公司设计了其自有的 SSD 驱动器，其称为企业级闪存模块（EFMs）。这些定制化的 eMLC 驱动器工作在 Nimbus 的 S 系列和 C 系列存储阵列中。EFMs 有专用的控制器软件和额外的本地非易失性 RAM（NVRAM），使得写操作完成后迅速确认而无需等数据回复到闪存上。

Violin Memory 公司使用其定制化的 Violin 智能化内存模块（VIMMs）。这些驱动器在三口的交换机中整合了闪存和控制器逻辑，提供一个阵列内和其它 VIMMs 的冗余连接以确保可用性。

(来源：TechTarget 中国 作者：Chris Evans 译者：张瀚文)

全闪存阵列解析：SSD 设计及全闪存应用案例

SSD 的设计

如上文所述，一些阵列供应商开始在其传统的存储阵列产品线中提供全 SSD 的产品。随着这些设备较使用磁盘驱动器的更为快速，但较之于为固态驱动器定制的阵列而言无法实现那种惊人的性能提升。传统存储阵列的设计中默认磁盘驱动器是整个基础架构中最为缓慢的部分，因此这种设计在阵列级的闪存中进行复杂的算法来预先从旋转的磁盘中获取数据。而这种特性在应用 SSD 时没有带来丝毫益处；SSD 所带来的数量级的性能提升需要全新的基础架构设计。

一个很好的实例包括从 SSD 阵列控制器中去除缓存部分。Whiptail 技术公司在其控制器中不使用任何缓存，而 Pure Storage 公司，另一家全闪存阵列供应商，将闪存部署在磁盘架上。Nimbus 和 Violin 则将 NVRAM 闪存替换成闪存模块。去除控制器中的闪存意味着在各控制器的高可用性配置中不需要任何额外的一致性逻辑控制。这同时简化了设计并提升了性能。

新兴企业占据了全闪存阵列市场

高端的全闪存阵列市场的主流供应商有 Kaminario 公司，Texas Memory Systems 公司，Violin Memory 公司和 Whiptail Technologies 公司。其价格在 15 美元/GB 到 20 美元/GB 之间，单台系统可提供高达 1,000,000 的 IOPS。在中端，Pure Storage 公司和 SolidFire 提供的产品在 100,000 IOPS，价格大约在 5 美元/GB 到 10 美元/GB 之间，使用数据精简技术来提供更低的价格指标。

一直到最近,传统的存储阵列才开始使用全 SAS 背板连接。而对于所有的磁盘驱动器而言,快速的 6Gbps 并不必要,因为传统的磁盘驱动器并不能充分利用 SAS 连接。而另一方面,SSD 则能够轻易地利用起 SAS 的全部速度,因此全 SSD 阵列需要完全的 SAS 带宽连接到所有的后端设备。

除了在 SSD 中使用损耗平衡技术来确保所有内存单元的均衡磨损,并延长有效使用时期,全 SSD 阵列的供应商还通过阵列级别的损耗平衡技术,来确保所有 SSD 和闪存模块上的 I/O 均衡,优化 IOPS 并进一步延长闪存的使用周期。

许多供应商同时还利用压缩和数据重复删除技术来增加其阵列的存储效能。重复删除在 SSD 上工作的更好,因为重复删除所带来的随机 I/O 问题不会给 SSD 造成性能延迟。举例来说,Pure Storage 就在其产品中使用大量的数据精简技术,以及可变的 RAID 条带化技术和动态 RAID 重组技术来提供连续的 I/O 性能。

全闪存阵列的应用案例

近年来普遍的认识是通过分层数据技术将数据部署在不同的存储层上,这种方式 and 全 SSD 的阵列似乎并不一致。不过,分层技术的出现是因为存储阵列的发展以及需要应对各种不同用途的数据。固态存储阵列并不能实现传统阵列那样的高扩展性,并且很明显的是相对昂贵的 SSD 使得全闪存阵列只适用于高 IOPS 应用环境,而非通用存储。

通常来看,SSD 阵列定位并适用于一致性和低延迟要求较高的应用环境,这包括数据库、虚拟桌面、游戏网站以及金融交易系统。

我们已经在高端应用中看到一些 SSD 阵列供应商,其提供 500,000 到 1,000,000 的 IOPS 以及低于毫秒级的响应时间。这些供应商有 Kaminario、Texas Memory Systems、Violin Memory 和 Whiptail。价格通常从 15 美元/GB 到 20 美元/GB。

主流的阵列供应商,诸如 Pure Storage 和 SolidFire 公司提供性能在 100,000 的 IOPS,以及 1 到 5 毫秒的响应时间的产品。价格较这些高端产品略低,在大约 5 美元/GB 到 10 美元/GB 之间。

新的参数

作为衡量存储阵列的成本,每 GB 美元这一参数使用了许多年,尤其是随着磁盘驱动器价格的不断下降以及驱动器容量的持续上升。然而,每 GB 美元并不能有效衡量高端的全闪存阵列,其定位主要并非提供通常的,集中化的存储,而是满足特定的应用需求。这时延迟和 IOPS 变为关键,以下两种指标更为适合:

- 单位美元的 IOPS，用以衡量产品是否货真价实
- 单位 GB 的 IOPS，用以衡量存储阵列的 I/O 密度

中端供应商，比如 Pure Storage 和 SolidFire 选择通过数据精简技术，诸如压缩和重复数据删除技术来提升其 SSD 阵列的空间利用率，并且定位于和传统阵列竞争每 GB 美元。不过并不是所有的数据压缩技术都是一样的；虚拟化服务器和虚拟化桌面的数据压缩率较高，不过数据库中的数据就不是如此，因此这些供应商提出的每 GB 美元数值仅能用于参考，而非实际应用。

绝非昙花一现

所有的闪存存储阵列都为关键应用提供了很高的 IOPS 数值。我们将会看到整个市场格局根据 IOPS 和价格划分成主流产品和高端产品。虽然并不适用于全部应用，但全闪存阵列一定能在绝大多数企业级环境中找到一席之地。

(来源：TechTarget 中国 作者：Chris Evans 译者：张瀚文)

SNW : SSD 用户探讨混合阵列及全闪存阵列

多少闪存才足够？关于使用全闪存阵列以及混合固态驱动器和磁盘驱动器分别能带来的收益，SNW（存储网络世界）在周二为固态存储使用企业的存储管理员安排了专题讨论。

全闪存阵列的两位使用者和混合式系统的两位使用者参加了这次讨论。这四位都表示两种方式都有其优势，不过他们根据各自特定的需要选择了不同的方案。并且这四位都是因为性能问题才选择闪存阵列的。

货运汽车制造商，TTX 公司的业务解决方案工程经理 Stephan Hundley 表示其从 Whiptail Invicta 采购的 17TB 闪存阵列要比之前使用的 EMC Clariion，一种由大量磁盘驱动器和一部分固态驱动器组成的混合式存储，要便宜许多。

“传统的阵列中有不少性能和容量问题，当然还有预算方面的问题，” Hundley 说，“我为 Whiptail 部署这种更低价的解决方案，使用 100% 的 eMLC 闪存，即企业级多层存储。这种技术更具成本效益，能够更快地部署，并且我们在数据库性能方面看到了惊人的提升。而这在过去一直是企业的一块心病。”

Ron Stinson 则是 Rainmaker 动画娱乐工作室的 IT 和运营总监，他表示为了企业大量的渲染工作，他需要大幅的性能提升。而购买一台全闪存的 Alacritech ANX 网络附加存储阵列的方案，较升级其原有的日立数据系统公司的 BlueArc 磁盘 NAS 要便宜许多。他说闪存技术协助其在 30 天的间隔内满足其公司负责渲染的 500 台主机和 650 台工作站的两项渲染工程。Rainmaker 仍保留了其 HDS 的 NAS 产品，不过只是增加了闪存而非更新整个磁盘系统。

“我们预见到所有应用程序的加载时间将从 20 秒钟上升到 2 分钟，” Stinson 说，“我们在想‘我们是否需要另外一个 NAS 机头？’这确实能够提升性能，不过我们对是否能获取到所需的性能表示怀疑。”

在面临“海量渲染风暴”时公司的服务器数量将超过 1,000 台，Stinson 采购了 Alacritech 闪存系统并即刻感受到了性能提升。他表示在部署闪存之后其处理器利用率从 70% 下降到 40%，而 NAS 系统的网络负荷下降了 20%。同时系统的成本较另外增加一个 NAS 控制器要低 40%。

半导体制造商 Linear 科技公司的 IT 经理 Amit Patel，以及金融服务机构 Calypso 科技的技术营运总监 Brad Taylor，说他们需要较传统的磁盘阵列更多的性能，不过他们发现全闪存阵列太过昂贵。他们都选择了 Nimble 存储的系统，其包括 4 个 SSD 盘柜和 12 个磁盘驱动器盘柜，并且使用 SSD 作为缓存。

Patel 说起 Nimble 混合闪存给他足够的性能来运行 30 个虚拟服务器以及 Oracle 和 Microsoft 的 SQL 数据库。“Nimble 将绝大多数热点随机读取转移到闪存上，并将顺序写操作转移到磁盘上，因此这样更为快速，并且非常可靠，”他说，“即便在写入到磁盘的过程中，他们会进行数据压缩。当我开始听到这点时，我感觉这会损失性能，不过其实并非如此。”

Taylor 则表示 Nimble 代表了“一种恰到好处的适中方案——从闪存中获取到极大的性能，并且合理地使用 SSD。至于闪存，我们已经谈到他所能带来的巨大的，10 倍甚至在一些情况下达到 100 倍的，性能提升。在业务实际需要时，完全的闪存确实是一种解决方案。但我们无法让 Nimble 承担其无法处理的工作负载。这种技术使用磁盘存储并通过 SSD 进行扩展。”

PCIe 闪存仍然价格不菲？

两方面都谈到他们评估了服务器内部的 PCI Express (PCIe) 闪存卡，不过仍认为更适合将闪存放在阵列中。Stinson 表示如果价格合适，PCIe 闪存确实合适他的应用。“我们正在试图在工作站中使用 Fusion-io 的 PCIe 卡，”他说，“这是一项代价高昂的解决方案，不过当价格随着时间下降，这无疑是一项可选方案。不过现在并不是。相较于 Fusion-io，我可以以十分之一的成本并且获取到 45% 的效果。”

Hundley 说他已经认识到基于服务器闪存的价值，不过这并不符合他的需要。他说他并不想要将 PCIe 卡放在所有希望性能提升的主机上。“我们能够感受到这种方案的价值，不过我们更需要一个 SAN 共享的基础架构，并非是单点的，”他说，“我们完全进行了 Vmware 的虚拟化，并且我们必须提供所有的 ESX 主机。”

闪存对于应用程序性能的价值

Linear 科技的 Patel 说到使用 Nimble 存储的最大收益在于降低了夜间备份所需时间并且加快了从他数据库中产出报告的时间。“我甚至不认为我用到了最大的性能”他说，“我将 SQL 放在上面，还有 Oracle，并且产生大量的报告，而系统并没有变慢。我仍有大量空间用于业务成长。这种性能是惊人的，我 SQL 的同仁总是要求更多的报告。这种报告通常要花上 24 小时，而现在在早上他们的报告就已经出来了。”

Calypso 的 Taylor 说他最大获益在于能够生成更多交易。“在我们这行，每秒交易数量是一项关键指标，”他说，“我们看到 10 倍，甚至在一些情况下有 20 倍的交易量提升。”

Stinson说到Rainmaker的获益在于能够在同时提供更多的渲染项目服务。“现在我可以承接更大的工程，”他说，“我可以增加我公司的规模，达到数千台服务器。”

(来源：TechTarget 中国 作者：Dave Raffo 译者：张瀚文)

只要使用正确 混合式和全 SSD 阵列都可以节省费用

在部署固态存储时，价格往往是用户的一大顾虑，不过有些机构发现闪存系统实际上可以为一些特定的应用节省费用。

尤其是在有些情况下，专为固态技术所设计的较新的存储系统能够降低成本。美国新罕布什尔州的科尔比学院采用了 NexGen 的混合 SSD 阵列建设虚拟桌面基础架构（VDI），其成本只有传统磁盘存储的三分之一。而位于德克萨斯州的航空航天光学公司采购了 Nimbus Data 的 S 系列全 SSD 阵列，价格较上一家混合阵列供应商的方案低了大约 10%。

用于 VDI 存储的混合式闪存

科尔比学院的 IT 总监助理 David Blaisdell 表示他预计 NexGen n5 存储系统可以使其以 28 美元的成本部署每台 VDI 的存储。该推测数据基于在单台存储系统上可以支持 600 台虚拟桌面计算得出。该学院目前只有不到三分之一的虚拟桌面，不过 Blaisdell 表示其测试表明采用 SSD 技术的 NexGen 有足够的 capacity 支持 600 个点的桌面。

科尔比学院的桌面虚拟化项目一期包含了 200 台桌面。该学院采购了 300 个 VMware View 的许可证，并且 Blaisdell 说他的服务器硬件可以容纳 300 台虚拟桌面，支持最终计划的一半数量。

“我们有足够的性能应对 300 台桌面，并且我认为一台存储单元就可以有足够的性能和容量来满足目标 600 台到 700 台桌面” Blaisdell 说，“我们发现创建一台新的桌面时的峰值 IOPS 大约为 4,000。不过在测试 Iometer 和 Windows 7 时，我们可以为每台虚拟桌面提供 25,000 IOPS。如果可以在一台存储系统中容纳 600 台虚拟桌面，我认为是非常让人满意的。而且基于现有的数据来看，我们认为是可以达到这个目标的。”

Blaisdell 说性能是 VDI 存储部署过程中的最大挑战“我们并不需要那种能够容纳 TB 级别存储容量的大型存储，但我们希望有足够的性能。”

科尔比学院的主要存储是 HP 公司的 LeftHand SAN。Blaisdell 说他有考虑过增加更多的 HP 存储来启动第一期的 VDI，不过“这样我们将不得不回到一年前的状态，然后随着系统扩展不断增加存储。因此，这时我们开始考虑下一代存储供应商所能提供的高性能解决方案。”

Blaisdell 意识到应当使用闪存。NexGen 使用了 Fusion-io 公司的 IoDrive 固态驱动器，并且和传统的磁盘配置成混合模式。他在使用 LeftHand 时就认识一些 NexGen 的人，因此他引入了一台配置 1.28TB 固态存储容量和 22TB 磁盘驱动器容量的 n5-100 用于测试。在一个月以后，也就是五月份，他采购了这台设备。

“假如我们使用传统的磁盘来满足 600 台桌面，我们大概需要花费现在预算的三倍” Blaisdell 说，“我们可能只能够在第一期购买 200 台桌面，并在后期逐步增加以满足性能所需。”

NexGen 设备剩余的容量也仅仅用于 VDI，尽管 Blaisdell 表示他可能会在厂商提供异步复制功能以后，增加另外一台系统用于灾难恢复。他最终可能在 SQL Server 和其它应用上使用 NexGen 以满足性能所需，他说。

全 SSD 阵列加速 SQL 数据库

Michael Perry 是航空航天光学公司的信息系统副经理，他表示他发现 Nimbus 相较于主流存储供应商是一种更便宜的方案。在使用 Nimbus 之前，他说他曾考虑过采购 EMC 公司的混合存储，不过 VNX 仅配置不到半 TB 的固态驱动器就比他现在 S 系列的全闪存阵列价格高出了十倍以上。另一种选择是在目前表现不佳的 NetApp FAS2020 阵列上增加闪存，不过 2TB 闪存的价格大约要 100,000 美金。

“我们是一家小公司。我只是不想再向总经理汇报说‘你应该知道我已经花了 40,000 美元了吧？我并没有很好地利用这笔钱。因此现在我要在花 100,000 美元’” Perry 说，“EMC 不到半 TB 就要 64,000 美元。而 Nimbus 的 2.5TB 系统只要 37,500 美元。一台 EMC 系统的价格抵得上 10 台 Nimbus 系统了。”

Perry 表示他从去年 11 月开始考虑在另一个虚拟化大量航空航天服务器的项目中使用全 SSD 阵列。他在四月份一次展会上和 Nimbus 高层交谈并进行了大量调研后选定了这家公司。“我的技术顾问在展会后和他们的 CEO 取得联系，并谈论了相关体系架构。我们的问题是，‘是需要一个从头到脚的 SAN 架构，还是一个基于 PC 的驱动器架构？’”

最终让人满意的 S 系列是完全的 SAN 架构，Perry 安装了一台 Nimbus 系统，作为航空航天光学公司经常访问的企业资源规划（ERP）应用的存储，其使用大型的 SQL Server 数据库。他表示闪存阵列使其 SQL 备份时间从 18 分钟缩短至 3 分钟。

Perry 表示他和他的用户现在非常信赖闪存。他将原来的 FAS2020 磁盘系统用于备份，将一台放置在托管中心，并使用 SnapMirror 将服务器镜像迁移到离线站点。

“我就像是机车上的孩子，发现原来可以每小时走 150 英里，就再也不愿意回到每小时 55 英里的方式上去了” Perry 说，“一旦你体验到了 SSD 的速度和响应能力，你就再也不愿意回到过去那种方式了。我的用户也有他们自己的方式，他们希望快速接受到响应。”

Perry 有 24 位产线员工，他们使用 ERO 服务器用于培训报告、视频和其它高分辨率的内容，他说。“如果系统无法及时响应，你就得让 24 个人每分钟等上五秒。不久之后，每年损失的生产力就达到了 100,000 美金。”

Perry 将很有可能采购第二台 S 系列系统，他说。“在 IT 业有一种说法，你永远不能有太大的带宽、太宽的屏幕或过量的存储。否则你将快速地消耗掉你的存储。”

[\(来源：TechTarget 中国 作者：Dave Raffo 译者：张瀚文\)](#)

本期电子书由 TechTarget 出品

