



IT Knowledge • Business Results



集群存储： 今天的价值，明天的首选

迈向虚拟化数据中心最为关键的一步

Tony Asaro
高级分析师

2007年3月

目录

目录.....	1
介绍.....	2
使用人员和管理人员.....	3
症结所在, 流动的数据中心、技术的飞跃.....	4
真实集群存储.....	5
真实集群与管理集群.....	6
真实集群与全局命名空间.....	6
集群存储系统所面临的挑战.....	7
今天的价值, 明天的首选.....	8
ESG 的观点.....	10

介绍

存储系统本身的复杂性是管理存储环境最大症结所在。很多先进的存储系统复杂程度极高, 需要丰富实践经验管理者和持续的人工管理才能实现最佳操作。这都是由于这些存储系统虚拟化程度不高导致的。

存储系统虚拟化程度越高, 其管理就越容易, 越能够达到最佳操作状态。任何时候, 采用人工管理, 不论管理员是否有丰富的经验和技能, 都会使整个操作流程效率降低, 出错几率也大幅增加。

虚拟化可以实现在逻辑上通观和对实物资产进行控制。RAID 是最普遍的存储虚拟化模式之一。使用 RAID 的存储系统, 将数据分布在一组驱动器上, 为了预防某个驱动器出现故障, 将数据同时在一个或多个驱动器上进行复制。在其某些格式中, RAID 主要有两个作用: 提高性能和提供数据保护。虽然 RAID 在整个存储系统技术购置中是极为宝贵的资源, 它提供的虚拟化层次并不高。RAID 在存储虚拟化上只是个开篇。

虚拟化和集群之间又存在什么关系呢? 两者有着千丝万缕的联系。集群是目前性能最强大最实用的一种虚拟化形式。在一个集群里, 多个物理系统可视为一个单个逻辑系统(与虚拟化相似)。因此, 集群存储不受单个物理设施的限制, 能够有效的消除“BOX”的界限。这些限制一旦被消除, 存储系统就成为一个“无限”的系统, IT 部门再也不用局限于做那些日常的琐事(如管理实物资产等), 能够发挥其更大的作用, 从而在公司整个业务中占据重要战略地位。

大多数情况下, IT 部门通常只是公司内一个一般管理部门, 作为业务的一个附庸, 更为不幸的是, IT 通常不被视为业务发展战略之中。有些时候, IT 部门最大的“敌人”反而是它自己。在业务部门的概念里, IT 部门的是一群聪明人做一些简单的工作, 如按按开关, 调调设备等, 业务部门通常都认为这就是 IT 部门的价值所在。但是 IT 部门认为, 在经历一些事情之后, 这种观点肯定会得到纠正。作为一名 IT 专家, 拥有别人所没有的丰富学识, 这是一件令人愉悦的事情。而且你的工作为公司作出了很大的贡献, 这给了你一种工作的安全感和尊重感, 人们对你的工作很放心, 而你也对每天以及今后的工作都有一个清晰的安排。不论事情难度和复杂程度与否, 你都能游刃有余; 工作越多越好; 那么最终你会成为 IT 的主管, 你的个人价值得到充分地体现。但是实际上, 这是一种错误衡量工作质量的方法。IT 部门应该是业务中非常重要的一个组成部分, 在业务中 IT 部门能运用技术强化竞争优势来为公司开发更新更好的产品和服务。

那些将 IT 列为战略部门的公司, 最终在业务上取得了成功。这个成功并不是做几个决策或执行几项政策就能够取得的。Benjamin Franklin 曾经说过, 疯狂的定义是 - 反复做同样的事, 却希望取得不同的结果。我不认为在数据中心这件事上我们作了一个疯狂的选择。然而, 如果我们耗巨资去采购和维护一个规模庞大、复杂的、半虚拟化的存储系统, 那么别指望在短期内能够解决我们的问题。

使用人员和管理人员

有两类人要受到我们的物理基础架构限制的影响 - 使用人员和被半虚拟化系统所局限的管理人员。

对用户而言,需要系统工作。他们能够运用各种软件,不受到技术的限制,自由读取数据。这也就是 IT 部门的矛盾之处。技术是用来提高我们的经验和效率的,但是一旦技术不能正常运行,过于依赖技术的我们,工作就很有可能陷入瘫痪状态。从这一点来看,IT 就成为了一项问题而不是解决问题的方案。IT 试图使其服务和技术对终端用户而言是隐形的,但事实并非如此。用户们非常了解 IT 人员及其局限性。

IT 部门经理受到各种技术、设施、工具、管理软件以及系统等的困扰,系统环境越复杂,IT 部门就越难满足业务和用户的需求。一旦物理存储界限和所属性发生变化,IT 资产会变得很瞩目。问题也就从这里开始了。

最常见应付改变管理的方法之一就是保持原样。这也是为什么我们通常总是购买容量超出目前需求的原因。我们创建一个较大的卷,需要占用大量分配容量,但是却很少使用。我们为了将来少做变动,总是在一开始就把什么都配置起来。我们预先购买的总是远远超过需求,这样就不用担心未来容量不足的各种后果,但是这样就导致资源的浪费,也带来出很多不必要的麻烦。当新的软件在线运行时,我们将系统容量扩大,购置新的存储系统和增加一些物理基础架构。当我们需要整合这些系统时,问题更为凸显,需要采购更新的更大的系统。然后我们被迫作出相应计划,实施从旧到新缩减存储资源。这样,就白白浪费了宝贵的时间、金钱和资源。

症结所在、流动数据中心以及技术飞跃

最终, 存储管理软件根本不能解决问题。GUIs 和相关软件的开发掩盖存储系统的复杂性, 但这并不是好的解决措施。这就好像在一个不完善的地基上建造房屋一样, 灾难是不可避免的。这并不是说好的 GUIs 和存储管理软件没有价值, 实际上, GUIs 和管理软件确实也发挥了其应有的作用。但是, 解决存储系统复杂性的关键在于存储系统基础架构本身。每个存储系统的复杂性才是问题的根源。

运用虚拟化将各个物理部件组合在一起, 在过去确实提高了效率, 如 RAID 和集群这个例子。在同时采用 RAID 和集群时, 多个分散的部件可被视为一个单个逻辑整体, 从而提高系统的性能和可靠性。另外一个关于虚拟化的例子更具有普遍意义, 即虚拟服务器的运用。终端用户在单个物理设备上建立多个虚拟服务器。尽管集群和虚拟服务器从某些方面作用是相反的, 但是它们的最终目标是一致的: 更好的利用物理 IT 资产来提高效率。

关于流动数据中心这个概念已经讨论了很多年。集群是流动数据中心的具体体现。一个真实的集群存储系统有自己的网络, 在其所有物理设备之间共享通信和元数据, 从而构建出一个高虚拟化的系统, 这是最根本的一点。如今, 网络与存储之间实际上是相互矛盾的, 存储系统只实现了部分网络化。我们做了大量的工作, 建立部分网络化, 使主机系统能共享存储系统资源。但是, 存储系统本身没有可共享通信网络。这些分散、独立的设施是通过网络连接到主机系统的, 他们自身之间不能够连接。因此, 从真正意义上讲, 网络存储系统并没有实现真正的网络化。

为什么我们要实现存储系统完全网络化呢? 将所有 IT 设备网络化的价值是显而易见的: 在这样一个存储环境下, 我们能够使用具有扩展性的系统, 随时按需添加资源, 你可以随时对性能、容量、连接性等进行升级, 数据能够在不同层之间迁移, 从而达到投资回报最大化。当你的存储环境增长时, 管理仍然很轻松。你可以扩容和执行负载平衡。根据需要对性能升级, 同时, 可靠性也随之提高。管理 1PB 容量的系统就像管理一个只有 1 TB 容量的系统一样简单。尽管物理基础架构中存在硬件设施增加, 有各种硬件, 作为一个大型的并行逻辑系统仍然便于管理。

只管理一个系统相对而言很轻松, 但是同时管理很多系统就具有一定的难度。人不具有可扩展性。因此, 我们就要对系统进行扩展。用更多的管理人员支持 IT 环境收效甚微。将多个系统作为一个逻辑部分进行管理是扩展 IT 管理能力的出路。集群能够让有效管理多个物理系统, 就像管理单个逻辑系统一样。

一个真实集群存储系统从逻辑上而言就是一个系统。过去, 一个人可以管理数 TB 的数据, 现在却可以管理数 PB 数据, 这样在效率上有了很大的跨越。

真实的集群存储系统

N-way 存储架构支持单个集群中多个存储控制器, 尽管这些集群中包括多个存储控制器, 但是在应用软件和用户看来, 就好像只是一个逻辑系统一样。集群是强有力虚拟化方式之一, 能够将多个物理设备组合成一个逻辑系统。在集群存储系统里, 用户可透明的添加 CPU、内存和带宽, 用户可基于业务的需要随时进行扩容, 而不需要专门去采购一整套新的存储系统。在传统的具有固定构架的双节点存储系统中, 当用户环境的增长已经超出其存储系统容量后, 用户不得不另外购置系统提高性能和延展性。集群架构可以将所有硬件资源进行集合, 根据需求, 线性升级性能。

集群存储系统能够开始以小容量的方式起用, 然后可根据需要升级到大的并行系统。对用户而言, 这是一种非常经济有效的方法, 不用提前购买超过其需求的大容量的系统。这也是传统固定架构的存储系统所无法比拟的。此外, 用户若想持续提高性能, 只要向集群添加更多的存储控制器即可。N - WAY 集群存储系统能够随时进行扩容, 操作简便。因此, 管理一个有 100 个节点的集群存储系统就跟管理只有 1 至 2 个节点的存储系统一样简单。虽然该系统包括更多的物理存储控制器, 但是仍然可视为一个单个的逻辑系统。

不是所有的集群存储系统都能提供同等性能。ESG 中所提及“真实集群”具有上述性能。真实集群存储是一个单个逻辑存储系统, 任何应用软件都能够通过集群的任一存储控制器读取存储其中的数据。一个真实集群能够将所有的硬件资源集合到一个单一系统中。

真实集群是一个高度虚拟化的系统, 每个存储控制器节点都是与其他节点相互沟通的, 共享一个“大脑”。这样导致的结果就是主机端口、CPU、内存、内部带宽以及容量都是在物理资源的逻辑池中。多个 CPU 并发工作, 来运行多个系统和进行 I/O 操作: 缓冲存储池能够扩展到数百 GB 容量, 主机端口构建出一个虚拟的通道与几十倍的 GB 带宽连接, 带有数百甚至上千个驱动器的磁盘将同时进行数据的读取和写入。

处理读/写请求的操作、进行当地和远程数据复制以及执行 RAID 的重建都是通过一个资源池来进行的, 该资源池可以按要求随时进行增加。最为重要的一点就是: 真实集群存储系统最终实现并行操作, 容量可以自由扩展。传统的企业级存储系统也能够实现并行操作, 但是前提是你就必须购置安装一整套系统。这在开始就需要大量的投资, 维持 3 - 5 年的需求空间。而采用真实集群存储系统, 则随时可以按需添加资源。

如何实现能简单的按需增加资源很重要。一个高度虚拟化的系统能够支持真实集群架构的前提是: 无论集群可以增长多庞大, 系统初始化容易, 增加资源过程简单, 逻辑存储系统在整个生命周期易于管理。

扩展性是相对而言的。集群存储并不仅仅适用于大型的 IT 环境, 当某个 IT 人员处理 PCs、服务器、网络、网络连接、安全、应用软件支持、服务器和存储时, 都将面临扩展性的问题。从另一方面来看, 在大规模环境下, 拥有更多资源, 就意味着可能要管理更多的基础架构和满足更多使用者期望。拥有轻松管理的存储系统让每个公司和机构都受益, 不论其规模大小。

采用真实集群存储还有一个重要的优势就是其前瞻性。采用传统的双节点存储系统,一旦某个控制器出现故障,那么包括系统处理器、缓冲器和系统内存以及内部带宽在内的 50%的资源都不可用。而采用真实集群存储系统则不同。采用三节点存储集群,大概 67%的资源处于可用状态;采用四节点集群,大概 75%的资源处于可用状态,而五节点则是 80%,依此类推。真实集群存储系统能够让在即使控制节点失效时,程序在可接受程度内运行。既要考虑用户和应用程序在最佳环境下运行,也要考虑到在不利的坏境需求,这是非常重要。作为平衡点,真实集群存储系统不需要过多配备工程师,你可以根据你的需要进行选择。此外,真实集群存储系统通常选用通用硬件,成本不会成为沉重的包袱,构建一个容量更大的集群不需很高成本。

如今,我们采购电力资源、内存、容量和带宽往往超出我们的需求,通常我们都是按照峰值加一点余量来进行规划,以满足未来 3 - 5 年的需求。结果,我们浪费了大量的金钱。每年,CPU、内存、容量和带宽的价格都在不断降低,提前购买这些资源就是浪费金钱,特别是在某些情况下,我们根本没有充分利用这些资源。磁盘容量的浪费是一个很普遍的现象。几乎每个用户都会告诉你,给他配置了某些资源,从来没有使用过。在某些情况下,大约 50%容量从未用过。此外,我们也很少提及浪费掉的电力、内存或带宽。这些就更难衡量了,但是这些跟磁盘容量的浪费相比,可能更惊人。而采用真实集群存储,则可以避免这些浪费。

真实集群与管理集群

ESG 将另一种集群存储称之为“管理集群”,这种集群,管理简便,但是却没有扩展性,也不能进行自由读取。管理集群最普遍的措施就是将两个控制器节点与相互运行的集群结合起来,将这一对控制器节点与另一对控制器节点相连接。每对控制器节点代表一个逻辑系统,任何主机都可以通过任一节点读取数据。每对双节点代表一个逻辑集群,能够将两个节点的硬件资源集中到一个系统上。所有其他节点成对的运行,这样各对节点间就不会出现性能累积了。管理集群的目标是为了简化整个系统的管理。然而,与真实集群相比,管理集群没有线性扩展性能。

真实集群和全局命名空间

在 NAS 环境下,一项颇具价值的虚拟化技术 - 全局命名空间出现了。全局命名空间很容易和真实集群的概念相混淆,因为这两种虚拟化形式在某些方面是一样的。全局命名空间代表一个逻辑层,通过这个逻辑层,人们能够浏览、读取和控制多个文件系统。如果你有 2 个或更多的文件系统,全局命名空间技术就能够同时将其作为一个文件系统显示给用户和应用软件。全局命名空间的价值就在于简化多个文件系统的管理。在某些情况下,有些公司的 NAS 系统中包含几十上百甚至上千个小型文件系统(容量从 2TB 到 16TB 不等),每个都是进行单独管理的。采用全局命名空间技术,系统管理员就能够更加方便的管理大量文件系统和所有的客户端。此外,全局命名空间能够更容易对数据进行读取,因为只有一条逻辑通道而不是数十、上百或上千个。使用全局命名空间另外一个优点就是数据能够在各种文件系统之间透明的在线传输。全局命名空间还可以用来为现有的异构 NAS 解决方案提供辅助或是作为同构 NAS 产品的一部分。不论全局命名空间以何种形式呈现,使用它,还是能够让 NAS 方案更易于管理,并充分得到优化。

真实集群存储也有上述功能。真实集群存储系统属于新一代解决方案,设计初衷是保持永不会复杂化、永远保持优化性能。真实集群本级提供单一读取层。真实集群 NAS 系统代表一个单一网络驱动器或安装点,用来支持高容量,可存储大量文件。真实集群 SAN 系统也有同样的作用,但是通过提供虚拟卷来实现的。与全局命名空间一样,采用真实集群存储,由于其便于使用,系统和管理员以及用户都能从中受益。真实集群能够进行硬件资源的集合,这是全局命名空间所不具备的。

值得再次强调的是, 真实集群存储系统将所有的硬件包括 CPU、缓冲内存、带宽和容量都统一到一个存储池中, 构建一个大型的平行系统。所有数据都可以在上面运行, 不论这些数据存储在这个集群内的哪个地方。有些全局命名空间方案还可以通过手动或自动负载平衡, 来提供一定程度的性能扩展, 但是功能远不及真实集群。

值得关注的是, 真实集群存储系统和全局命名空间解决方案是互补的。在使用全局命名空间解决方案时, 你可以运用真实集群存储系统, 发挥其优势, 来运行一些新的软件, 同时, 用全局命名空间管理现行复杂的 NAS 环境。如: 你可以使用异构全局命名空间技术将数据从传统的 NAS 系统上迁移到一个下一代真实集群 NAS 解决方案中。在这种情况下, 全局命名空间技术执行一次性数据迁移, 或作为一个智能分级存储环境的一部分来使用。

集群存储系统所面临的问题

可以这样说, 我也是集群存储的支持者。当我们支持某个具体的假设并收集所有的证据来支持这个假设时, 就会认为一切都是及其合理的。为了确保我没有偏见, 讨论集群存储系统所面临的问题也是很关键的。

当真实集群存储技术大行其道的时候, 就它自身而言, 这还远远不够。我们必须接受和充分运用已经有的产品, 而不只是在简单的在技术层面推崇。当评估这些解决方案时, 首先审视产品由来, 评估其管理软件、特点、其他功能以及产品路线图等。不要仅仅了解这个产品本身, 其公司背景也要深入了解。了解该公司是否有好的合作伙伴, 是否能够提供优质的服务支持。

了解真实集群存储系统如何进行微量扩展的。某些真实集群存储系统能够一次扩展一个存储控制器节点, 而有些却要求多个存储节点以组添加到集群中。

大多数供应商支持真实 N - way 集群存储架构主要为实现性能可扩展性。确实, 这些存储系统有扩展性, 但是并不是所有类型的数据都能够进行扩展。有些系统杰出性能只适用于大数据块。一些系统高性能只适用与读数据, 而不适用与写的操作。还有一些只能对有续数据进行优化。很多集群存储系统采用普通的服务器和部件, 避免使用任何专用硬件, 以降低了成本。但用通用件达到专用件性能水平, 常常导致采用大量通用件, 造成资源浪费。虽然普通硬件的性价比更具吸引力, 你需要更多的物理系统来满足性能的要求。

在整个存储系统生命周期内降低管理难度比系统的延展性能更重要。即使在非性能密集的环境中, 在评估存储系统的时候也应该考虑集群存储。然而, 值得关注的是, 不论多大规模的集群, 真实集群存储系统都具有使用简单、维护方便的优点。

今天的价值, 明天的首选

与人们的通常观念相反, 我们不喜欢 IT 行业更新太快。事实上, 技术革新一直不受欢迎是因为它会干扰中断我们当前的环境。对于我们正在使用的环境, 我们投入了大量的时间、金钱和资源。对于现有复杂的系统管理以及围绕这些系统进行最优操作方面, 我们已经拥有丰富的经验。我们的自然反应就是继续投资更多的时间、金钱和资源来使用更多的同样的软件和硬件, 从而使生产效率达到最大化。

我把这种现象称之为“西西弗斯综合症”。“西西弗斯”就是指一个杰出的人, 一生困扰重复着做把山石推上山, 之后看山石滑下来这件事。即使在某种情况下, 我们取得一些进步。也许这个山石会变小或这座山会变得平缓一些, 但是这些都是暂时的。其他的因素一直都在变化, 会使得事情更加复杂和困难。我们每天都在努力将山石推上山。再看它回到起点。

福布斯榜上排名前 100 位的某家公司管理人员告诉我, 他的终极目标就是, 使存储系统达到用户要求, 并且易于使用。那些所谓的先进的存储系统都不能满足这个要求, 操作复杂、半虚拟化, 设计的初衷就不是便于用户使用。有趣的是, 这个系统管理员同时管理几十个复杂的传统存储系统, 并没有更换这些系统的意向。那么我猜想, 他肯定想将现有的系统简单化, 并且保留现在所拥有的存储系统。但是这是不可能发生的。对于他以及很多与他相同的人而言, 将所有系统都换掉是很不现实的。是吗?

现在, 传统的处理软件大多与传统的存储软件联系在一起。在一些大型公司里, 这仍然是现实, 并且不会有什么改变。然而, 大型的公司采用分级存储。基于磁盘存储归档的集群架构是使用最广泛的一种。某些集群存储单个系统支持容量可达数百 TB 甚至数 PB。

中型企业更倾向于将集群存储系统作为其主存储系统, 他们在现有架构基础上投资相对少一些。对新技术接受能力比大企业快。这些公司的存储容量可达 100TB 以上。因为他们资源较少, 所以他们需要一个易于管理的存储系统来支持业务运作。

现在我们使用存储的方式在发生变革。业务关键性应用远不止数据库、小型交易数据类和邮件。文件型在线应用软件, 多媒体型数据逐步被认定为业务关键数据, 产生收入的来源。某些情况下, 大量的用户下载和上传这类大型文件、图片、视频和音频文件, 如 iTunes, MySpace 以及柯达 EZShare 等。数百万用户在互联网上上传数据, 并下载其他内容。实际上, 我们已经意识到, 每天越来越多的人上传大量的数据(数以百万计的文件), 同时下载大量的数据。

业务随时在改变中。目前我们是采用一个存储系统去满足所有要求, 用同一存储系统存储所有数据, 这是不正确的。就好像我们只采用一个方法来解决所有问题一样。这样尽管做起来简单, 最终却导致很多问题, 对经济和运作造成消极的影响。

对于那些已经在解决方案上投以重资的企业, 建议您采取对新应用软件时评估高虚拟化存储系统。在不干扰现行工作的基础上引进新的技术。一旦你认为这些解决方案适合您的需求, 你就可以按需进行扩展了。最终, 实现用下一代高虚拟化的解决方案取代传统的复杂的部分虚拟的存储系统。享用高虚拟化的解决方案提供的整个存储生命周期的使用和管理简便性。

既然集群存储使用价值如此之高, 为什么没有得到广泛应用呢? 因为它需要时间来让人们了解。任何新技术或新架构的出现都要花费好几年的时间成为主流技术。举例而言, 汽车开始出现的时候, 大约花了 45 年的时间在美国市场占有率才达到 25%。在我们开发集群系统的时候, 我们面临了很多的问题、偏见、管理压力、谣言、困惑、投资时间、金钱、资源, 以及人们的不了解和误解。这就是任何新技术走向成熟都要面临的问题。

ESG 的观点

如今, 高度虚拟化的集群存储系统首次向人们传达这样的信息 - 它能够推动业务的开拓。然而, 更大的目标是将数据中心各级都虚拟化, 从而让资源充分流动起来。只有这样, 才能突破界限。

ESG 深信, 存储系统需要高度虚拟化。存储系统应该能够任一扩展, 而不是定容性质的。我们应该将存储系统的扩展从固定的、有限的程度发展到灵活的、无界限限制的程度。

从长远来看, 应该出现一个逻辑存储系统, 来满足我们所有的存储要求 - 业务关键型, 要求苛刻的应用软件到数字归档和第三方的应用软件。存储系统应该是高度虚拟化, 支持文件, 数据块, 多个物理组元, 分级存储, 支持智能数据管理和数据保护功能。这一前景只有通过高度虚拟化和集群构架能实现。再下一代存储系统应该将 IT 从位, 字节, 系统容量、带宽、处理动力, 等琐事中解放出来, 让他们集中精力发挥其业务战略作用。同时, 再下一代存储系统应该为用户提供最高水准的容量和可靠性高同时, 易于使用。

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources the Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482-0188.

Isilon Systems, Inc. | www.isilon.com
3101 Western Ave, Seattle, WA 98121

Toll-Free: 877-2-ISILON Phone: +1-206-315-7602
Fax: +1-206-315-7501 Email: sales@isilon.com

Isilon 中国
北京朝阳区工体北路甲2号盈科中心1441室, 100027

电话: 8610 65391262
传真: 8610 65391297