

# 大容量数据存储 解决方案

## 大容量数据存储现状

人类已经步入一个数字化的信息时代，IT在社会生活的各个领域中正处于前所未有的关键地位。IT部门的各种业务数据的数据量近年来呈几何级增加，这些数据必须存储很长的时间并且确保其可访问性。大量的业务数据保存在非常昂贵的第一层磁盘存储上不是一个明智的选择。虽然存储虚拟化技术有助于通过整合数据存储部分解决这一难题，但是，仔细分析他们的存储使用情况和性能需求，并实施分层的存储体系结构中，组织可以针对存储基础架构与业务需求寻求最佳解决方案。

## 新技术满足不断增长的存储需求

IT管理人员的最大挑战之一，是大量数据高效、安全地长期保存，快速和简单管理维护。当下IT管理人员需要关注的主要问题是需要存储的数据的爆炸式增长。根据"The Digital Universe"预测，"从现在到2020年，世界范围内创建和拷贝的数字信息的总量会难以置信的快速增长，各种媒体形式例如声音、电视、电台、电影、文档将不断从模拟存储转换到数字保存，数字宇宙将增长几乎快到120万PB的容量或1.2 ZB。这种爆炸式的增长意味着到2020年，我们的数据总量将比2009年增长44倍。短信、电子邮件、文档、图片、视频、社交网络，等等采用云存储服务和嵌入式的系统进行保存，他们将增长近67倍。虽然人在家里，但工作中还是在移动设备上进行，企业估计有70%的全数字化的信息内容存储在IT部门。因此，IT管理团队肩负着存储系统的基础设施建设以及数据安全保护的重任。然而，文件的增长速度在很大程度上增加了存储管理员本已沉重和复杂的工作量。据IDC预测，到2020年，全球IT专业存储管理人员的数量增长速度仅为1.4。人才资源紧缺的存储管理员要管理更多的数据，IT部门必须提高运营效率。提升存储效率明显的技术至今已经有云计算、虚拟化技术，它利用虚拟化整合工作负载，提高资源利用率，最大限度地减少昂贵的空闲时间。IT部门还可以动态地分配硬件资源时需要它们的地方，无需过度配置，以满足高峰负荷的需求，综合系统提高IT效率，减少了管理存储系统的维护时间。

## 五个海量存储的最佳实践

### 1. 最大限度地分配和利用

- 大多数存储系统都存在分配空间浪费和使用率浪费的现象。

实施存储资源管理，可以把浪费的容量恢复到工作状态，可以降低第一层存储成本。

### 2. 存储虚拟化

- 存储虚拟化，自动精简配置和存储资源池共享。这将显著降低空闲的和浪费的存储空间。

### 3. 采用分层存储

- 分层存储针对应用的容量、用途和性能需求，选择不同存储介质以降低总成本。

将驱动器分类并迁移到不同的存储层中，存储分层管理战略大大简化了存储资源管理。

### 4. 堆栈管理

- 冗余阵列中虚拟化管理程序和优异的存储资源管理器在服务器堆栈今天已经取得了先进的管理能力，管理复杂和昂贵的。从堆叠中的存储管理，大大简化管理和降低成本。

### 5. 使用企业级存储

- 存储是专门为堆栈的管理环境，消除冗余管理能力的成本。

这套存储系统必须是高性能，超可靠的，易于管理和高效运行。

虽然虚拟化的投入一直以来都专注在服务器上，最近几年，IT组织已经逐渐开始在存储虚拟化部分加大投入。存储虚拟化的优势在于将存储基础设施进行整合和优化，而且有可能降低近一半数据中心成本。一个新兴的存储虚拟化技术已经被证实特别有效，尤其是针对于堆栈（服务器）管理，它被应用到企业级存储系统并嵌入到SAN存储区域网中。

这种类型的存储管理方案利用内置的虚拟化管理程序的存储管理能力。这意味着，不需要额外购买包括重复的存储管理功能的存储系统，存储系统可以内置一套简单的软件，将大量的存储管理精力留给虚拟服务器堆栈管理。这使得管理员能够有效地管理更大的存储量和降低管理存储的成本。

但是，为了更好的使用存储虚拟化技术，管理员需要完全理解他们的应用环境的存储需求。特别是他们必须确定存储的数据的特性，参照读写行为和性能要求，使用这些信息来建立一个分类方案，定义不同价值的数据以及服务水平，以支持应用程序的性能。只有针对准确的数据存储和性能分类与使用的存储硬件的能力来建立IT基础设施，IT系统的效率才会提高，成本才会降低，大量的数据才会得到存储系统的长期保护。

| 存储保护层 | 等级分类                                                                                                      | 性能 |
|-------|-----------------------------------------------------------------------------------------------------------|----|
| 1     | 关键任务的数据 <ul style="list-style-type: none"><li>企业核心数据，必须保证可持续访问</li><li>最高的性能，零停机时间</li></ul>              | 最高 |
| 1-2   | 关键业务数据 <ul style="list-style-type: none"><li>非常重要的业务数据，需要经常访问</li><li>高性能，高可用性，需要4小时以内恢复</li></ul>        | 高  |
| 2     | 可访问的在线数据 <ul style="list-style-type: none"><li>企业必需的，不经常访问的，对成本敏感</li><li>在线高性能，高可用性，需要8个小时以内恢复</li></ul> | 中  |
|       | <ul style="list-style-type: none"><li>不变化的数据，备份恢复</li><li>- 非托管档案，对成本敏感的 • 磁盘性能，自动恢复</li></ul>            |    |

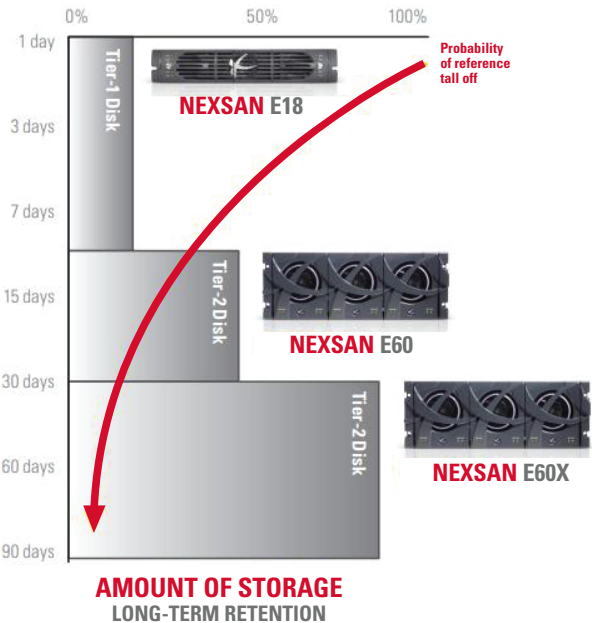
### 设计您的大容量存储系统架构

本文将侧重于阐述长期、大容量存储，因为它代表了绝大多数的数据，优化管理这种类型的存储对组织长期的存储需求的影响最大。

通常情况下，长期，大容量存储用于非结构化数据（如音频，视频，电子邮件或文件），这些数据占总存储容量的80%。组织也需要能在这个存储系统上进行数据的快速检索。因此，存储基础架构必须允许操作系统，应用程序或用户积极地引用这些数据。从长远来看，大容量存储不包括数据存储在数据库中的，面向对象的数据必须保留很长一段时间（或者将其归档以满足法规遵从）。

#### 解决方案

## TIERED STORAGE NEXSAN E60X



| 数据保存天数 | 访问概率  |
|--------|-------|
| 1      | 75%   |
| 3      | 50%   |
| 8      | 25%   |
| 30     | 2%    |
| 90     | 接近 0% |

### Nexsan E60X 分层存储

如何建立一个高效的海量数据存储库，优化利用堆栈管理的虚拟化存储？第一步是将支持的应用程序类型和非结构化数据进行存储库分类，建立存储的性能的服务水平。非结构化数据的性能要求是不断变化的。例如，一些企业可能会考虑，Exchange服务器是一个关键任务应用程序，需要非常高的性能服务水平。相反，文档可以具有一个非常低的访问概率，因此，需要一个低的服务水平的性能。分类数据，根据所需的性能水平，不仅有利于界定技术要求的基础设施，它也体现了最佳的存储分层管理方法。

第二个步骤，随着时间的推移，数据访问概率会降低吗？

这有助于确定如何判断存储的数据是否有正当理由的存在，随着时间的推移将数据迁移到不同的存储层。

通常情况下，大多数的用户数据的访问行为随时间推移急剧下降。数据保存一天后，该数据再次被访问的概率大约是75%，在90天之后，该数据再次被访问的概率接近于0。经过分析，企业可能会决定将一些数据迁移到二级存储，并保持其整个生命周期。其他数据可以持续保存在高性能的一级存储中，这是性能最好、安全性最高、价格最贵的存储层，将惰性数据迁移到二级存储（拥有极好的性能和可靠性，不到一级存储一半的成本）。

信息生命周期管理是一个术语，用来描述一个堆栈中的分层存储基础设施所管理的分级存储管理能力。分层存储架构允许企业根据他们的业务需求调整其存储硬件的性能和可用性。当IT部门不积极进行分类和管理他们的非结构化数据在存储系统中的位置，存储系统中就会产生相当大的不活动并且没有被读取的数据。存储6个月或以上惰性数据占据大量存储空间，产生了资源浪费。

### 解决方案

Imation Hong Kong Limited. 27/F, Yen Sheng Centre, 64 Hoi Yuen Road, Kwun Tong, Kowloon

Email: Nexsan.Info.Ap@imation.com | www.imation.com/nexsan 或 http://www.nexsan.com.cn

© Imation Corp. Nexsan, Nexsan 徽标、E 系列、FASTier、E-Centre 和 NestOS 是 Imation 公司的商标。其他所有商标均为其各自所有者的财产  
(修订时间: 27/07/15)

一般情况下，过多惰性数据占据一级存储时，他们可以迁移到二级存储进行长期保存。最大的问题是，存储管理员通常需要采用具有成本效益的分层存储架构需要通入大量的时间和精力，进行管理控制。好消息是，成本效益软件大大简化了这一过程。存储管理员可以使用迁移管理器中的应用，建立简单的分类规则来定位、迁移和管理数据，来实现分层存储。

## 最佳存储实践：使用NEXSAN实现分层存储

当企业对数据进行适当的分类，根据预期的使用情况和性能的要求，将数据存储在堆栈管理的虚拟化存储阵列进行规划，分层管理，这样可以提高存储管理效率，并满足他们的存储增长的需求。

但是，需要特别关注的是 - 越来越多的虚拟化技术需要提供更高的I / O性能。数据库和存储虚拟化系统，如赛门铁克的Storage Foundation的VMware vSphere或Microsoft Hyper-V，对I / O处理能力要求很高。幸运的是，这些虚拟化环境已经包含了管理I / O路径的能力，以保持较高的性能。Nexsan的存储专门为那些没有内置的I / O路径管理功能，又想搭建SAN环境或iSCSI架构的情景提供了便利。Nexsan公司提供模块化，支持堆栈管理、二层长期保存的大容量存储阵列。Nexsan公司的模块化技术，增加了可扩展的能力和性能。

Nexsan的E系列存储可配置4个8 Gb光纤通道和4个10GbE iSCSI端口，支持阵列SAS扩展连接，提升存储系统性能。支持控制器高可用性无缝切换，存储系统针对虚拟化存储特性进行了性能优化，使得Nexsan的一个完美的解决方案能够有效地存储和管理不断扩大的数字宇宙。Nexsan的阵列是高密度存储阵列，每机架单元可容纳更多的驱动器，同时，内置的AutoMAID能力有选择地降低功耗，AutoMAID节能技术最高可节省存储系统85%能耗，降低了存储系统散热量，降低了冷却设备的负荷。

Nexsan的E系列一直是专门为中端市场的设计，通过提供领先的解决方案，降低了大容量存储与管理数据的成本，提高了管理效率。例如，企业可以一开始使用Nexsan E18，拥有18个3TB硬盘，提供54TBs的原始容量在2U的机架空间。随着业务的发展，他们可以很容易地增加一个Nexsan E60X扩展存储单元，E60X在4U的机架空间可支持240TB的存储容量。

### 解决方案

Imation Hong Kong Limited. 27/F, Yen Sheng Centre, 64 Hoi Yuen Road, Kwun Tong, Kowloon

Email: Nexsan.Info.Ap@imation.com | [www.imation.com/nexsan](http://www.imation.com/nexsan) 或 <http://www.nexsan.com.cn>

© Imation Corp. Nexsan、Nexsan 徽标、E 系列、FASTier、E-Centre 和 NestOS 是 Imation 公司的商标。其他所有商标均为其各自所有者的财产  
(修订时间:27/07/15)

## 结论

需要存储的数据量呈几何级数增长，这种增长不断持续。厂家已推出相应策略，分别针对存储管理系统、存储容量、文件管理功能和性能几个方面，这在几年前是不可想象的。专门为企业设计的存储系统注重更大的容量，更高的性能、更高的密度、更好的管理功能，以满足存储数据的需要。企业会发现，面对复杂的存储和服务器基础设施的，越早规划存储分类方案，越早分析不断变化的数据并使用存储分层技术，他们就会越早获益。