

# 第5章 数据复制与容灾

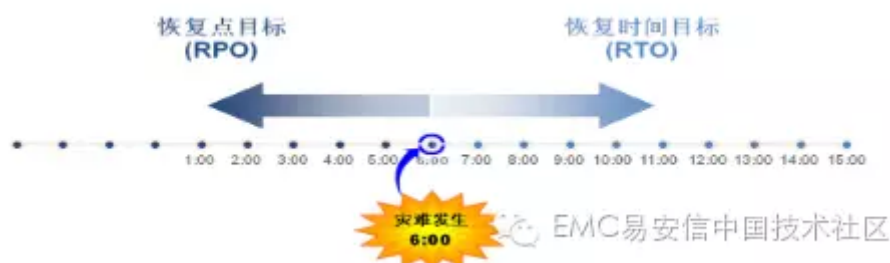
## 衡量容灾系统性能的两个关键技术指标-RTO与RPO

原创 EMC中文技术社区 [戴尔易安信技术支持](#) 2016-05-16

对于信息系统而言，容灾就是使信息系统具有应对一定的灾难袭击，保持系统或间断运行的能力。目前，大家比较习惯用一些技术指标来衡量容灾系统性能、需求等等，今天想借此机会介绍一下常常提到的两个关键指标：

**指标一，恢复时间目标 (RTO: Recovery Time Objective)**，其以应用为出发点，即应用的恢复时间目标，主要指的是所能容忍的应用停止服务的最长时间，也就是从灾难发生到业务系统恢复服务功能所需要的时间周期。RTO是反映业务恢复及时性的指标，表示业务从中断到恢复正常所需的时间。RTO的值越小，代表容灾系统的数据恢复能力越强；

**指标二，恢复点目标 (RPO: Recovery Point Objective)**，RPO是反映恢复数据完整性的指标，其以数据为出发点，主要指的是业务系统所能容忍的数据丢失量，见下图：



一般来说，RTO和RPO的值是根据实际的业务需求来确定的。从狭义上讲，容灾是通过在异地建立和维护一个备份存储系统，利用地理上的分离来保证系统和数据对灾难性事件的抵御能力。从广义上讲，任何提高系统可靠性与可用性的努力都可称之为容灾。

就EMC经典的容灾产品而言，试举两例：有SRDF(Symmetrix Remote Data Facility)，它可以提供范围广泛的远距离复制功能，见下图：



EMC的VPLEX也可以通过VPLEX Metro和VPLEX Geo等实现远距离的数据访问和复制，见下图：



## EMC备份数据到磁盘技术

戴尔易安信技术支持 2016-05-17

<https://v.qq.com/x/page/j0196aqtkc0.html>

21世纪最重要的企业资源就是数据。每家企业都把信息和数据视为自身发展和应对一切挑战的根基，对于数据的保存和备份被众多企业定为到了一个前所未有的战略高度。数据备份的重要性不言而喻。EMC作为全球范围内企业数据和存储方案提供商的领导者，正由它独特的备份到磁盘技术成功帮助无数企业实现关键业务。本视频将为大家阐述业务挑战及“备份到磁盘”技术如何设法解决，“备份到磁盘”技术的选项和益处，并简述信息生命周期管理及它与备份策略的关系。

# 数据中心变革 - 从主动/被动到双活

---

原创 EMC中文技术社区 [戴尔易安信技术支持](#) 2016-05-09

本文从IT服务提供商的角度，介绍了传统主动/被动数据中心在运营方面的劣势和过度到双活数据中心对业务的的优势，以及通过EMC VPLEX Metro实施双活数据中心的一些条件。

IT服务提供商在理想情况下定位为想客户提供解决方案，以便帮助客户具有灾难恢复（DR）需求的业务。可满足灾难恢复需求的可行选项围绕数据备份服务和将数据复制到异地数据中心发展。甚至在将辅助数据中心用作恢复站点时，它们仍在传统主动/被动实施中配置，而这仍意味着将给辅助站点的生产环境带来重大宕机。如果发生灾难，这些解决方案预计会产生宕机和业务影响。根据服务，恢复时间目标（RTO）可能是数天，而这意味着会给许多组织带来灾难性的业务损失。灾难恢复到辅助通常意味着在某个时间点上环境必须“回切”到原始生产站点，从而给业务带来更多宕机。

传统主动/被动数据中心，在传统主动/被动部署中考虑基于云的基础架构时，主站点可用性方面存在许多限制。例如，在多个阵列之间或在数据中心之间远程实施VMware环境时，会出现许多挑战。这些可能是通过VMware vMotion重定位工作负载方面的挑战，而此类重定位传统上仅在同一数据中心内才能实现，或者工作负载需要使用VMware High Availability（HA）在恢复数据中心重启时才需要。站点故障或重大基础架构停机会导致应用程序中断，以及在停机需要手动干预时进行故障切换，甚至在使用VMware Site Recovery Manager（SRM）等灾难恢复工具时（如下图所示）。可能会显著影响业务运营和收入。



VPLEX带来数据中心之间的真正双活部署。越来越多的企业对防止业务运行发生任何可能会导致宕机的时间方面感兴趣。EMC VPLEX Metro可简化这些部署模型且支持附加可用性，这对于总体体系结构至关重要。通过VPLEX Metro，允许您跨多个阵列扩展，VPLEX Metro不单单是将环境限制为单个数据中心，而是允许我们跨多个数据中心扩展环境，从而提供真正的双活环境。（如下图所示）。在图中的配置中，两个数据中心都运行主动工作负载并且服务客户应用程序。主数据中心的数据立即可用并与辅助站点同步。EMC VPLEX和VMware结合成为更加全面的解决方案，因此业务几乎或完全不会宕机。它们携手提供VMware FT，HA和DRS的所有自动化优势，同时允许跨数据中心实施，而不是限制在数据中心内部。



VPLEX Metro使用两个VPLEX集群（每个数据中心一个）并具有独特功能，可支持通过直写缓存存在两个集群间镜像数据的同步分布式卷。由于VPLEX Metro分布式卷有VPLEX缓存一致性算法控制，因此可从两个数据中心进行分布式卷的主动数据I/O访问。这让VPLEX Metro可提供可提供的远比传统主动/被动复制解决方案多的多，并且跨数据中心支持真正的双活基础架构。

VPLEX会将同一数据块卷分发到两个数据中心，并且确保标准HA集群环境（如VMware HA和FT）可使用此功能，因此也可在服务提供商数据中心之间透明部署。最起码的要求是VPLEX Metro让主机集群相信节点之间没有距离，因此它们的行为如同在单个数据中心那样。要实施此配置，需要满足两个条件。首先，对于VMware云基础架构，站点之间的网络延迟需要为10ms或更短，根据底层基础架构和服务提供商提供的网络资源，这通常会将此类部署限制在同一城区内。但是，服务提供商可能已有网络基础架构，允许他们在更远距离满足10ms的延时要求。关键是延时，不是距离。

## 满足变化中存储需求的高效架构

原创 EMC中文技术社区 [戴尔易安信技术支持](#) 2016-05-18

软件定义的存储的主要目标是将应用程序和数据与最适当的存储单元保持一致。而满足利用软件定义存储构建高效存储架构的关键是，从战略上制定政策来规划软件定义的存储解决方案需要整个存储基础架构的全面可见性。本文描述了一些实现软件定义存储高效架构的一些考虑。

构建高效的企业存储基础架构要求组织制定配置、部署和管理存储资源的策略。这超出了仅识别存储组件本身的范围，并需要深刻了解服务器、网络、软件层和工作负载与所有存储元素的关系。要建立支持显示所有管理元素关系的IT基础架构，了解软件的存储架构具有类似的属性。尤其硬件基础架构中抽象出池化资源，以最大限度提高应用程序和工作负载的可能性和性能。不过，对具有其他数据中心组件和存储资源实现池化要困难得多，这是因为底层硬件设备一般差异很大。而且，与其他设备不同，存储是所有企业数据的“家”，由于其庞大的规模和复杂性，本质上具有难以移动性。因此，软件定义的存储的主要目标是将应用程序和数据与最适当的存储单元保持一致。当然，每次调配新的IT服务时都要向存储系统手动分配工作负载会极其困难而且非常耗时，因此，自动化任何软件定义的存储实现的重要组成部分。建立策略实现存储资源调配的自动化，并确定适当的存储资源。这可以极大地缩短存储管理员花费在手动资源调配任务上的时间。

**从战略上制定政策来规划软件定义的存储解决方案**需要整个存储基础架构的全面可见性。应收集有关存储资源的配置，性能和容量的详细信息并将其存储在综合存储库中，而且需要采用一个集中控制台，在支持区提供所有存储元素的统一报告。这样，存储管理员只需查看单一界面来确定存储体系的体系结构和运行状况。应提供计算机所有层（如应用层、虚拟化层、管理层等）的视图，并以可视化方式映射各层之间的关系以便参考。这种整体的可见性可以帮助组织确定依赖关系，并确定一个区域中的条件如何影响其他区域。例如，可以很容易地识别多台主机争用共同存储资源，并且可以快速诊断和解决任何瓶颈问题。



**综合监控整个存储体系**（包括传统存储体系和软件定义的存储体系）无疑会产生大量的详细状态信息。这些原始形式的信息对于管理员来说只是“白噪音”，他们无暇阅读大量日志报告。利用分析来处理海量数据可以识别任何时间、潜在的问题或配置更改、并向支持人员报警，以防存储系统偏离既定的基准，并确保存储资源的高可用性。此外，这种主动监控以不间断的“闭环”形式全程监控任何新的IT部署和更改，确保它们的引入不会影响其他任何IT元素的性能。通过以历史趋势图的形式显示性能细节，还可以关联时间，并确定配置改进机会。例如，如果系统的IOPS在特定的时间大幅下降，此问题可能与同时发生的其他事件有关（如备份、修补或系统扫描）。同样，趋势报告中的表现出性能缓慢下降可能不足以启动报警，但可能表明需要额外容量或者跨存储单元优化负载平衡。

**整体可见性**对于实现主动容量规划也是至关重要的。通过了解当前有多少容量可用，谁在使用，以及预测何时需要更多容量，组织可以在预算限制范围内快速确定哪些特定资源需要扩展。在使用分层存储基础架构时尤其如此。因为每层都有自己的独特性能参数。一般来说，较低分层的存储更侧重于降低成本，而较高分层的存储则强调高可用性和性能。当每层接近其独有容量限制时，组织可以确定是否要扩展存储系统、减少工作负载和/或更改配置（如以性能为代价支持更高的密度）。采用精简资源调配的日益增加还要求密切关注容量可用性，因为它要求故意超额分配物理存储系统。尽管精简资源调配可以帮助组织最大限度提高其存储投资的价值，但会在很少或没有警告的情况下引起触发及硬性限制的危险。主动容量监控可以减轻这些风险，同时还实现了该方法的成本优势。

**集中存储管理**除了提供在增加存储容量时降低资本成本的必要见解外，还有助于减低日常运营开始。例如，消除一些常见任务 – 像执行数据中心审核，以确定存储资产状况和空间可用性。此外，及时发现问题和潜在问题再加上能够执行真正的根本原因分析，可以将存储管理从被动响应的做法推进到主动防范问题的流程，结束了系统性“救急”式的中断/修复循环。这可以减少支持工作，让管理员抽出更多精力投入到流程改进和以业务为重点的新项目中。这样，不需要额外员工支持环境增长，因此现有人员有更多的精力关注不断扩展的服务，而不是忙于维护日常操作。此外，使用自动化还可以简化管理流程，因此不太熟练的员工完全可以胜任更高级的任务。统一存储管理使企业能够最大限度地提高管理员的工作效率，同时提高实施绩效和可靠性。

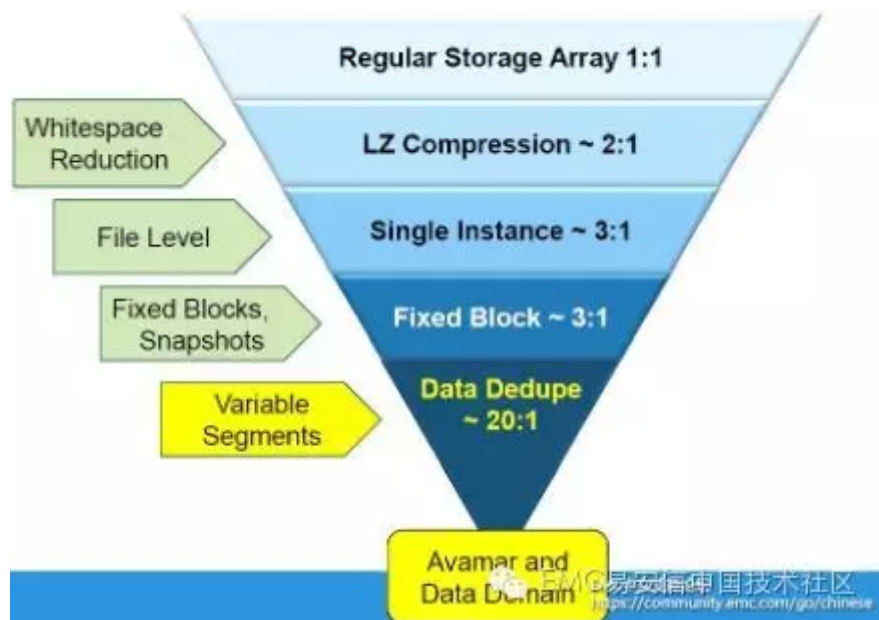
## 数据缩减技术效率对比

原创 EMC中文技术社区 [戴尔易安信技术支持](#) 2016-05-19

面对数据的急剧膨胀，企业需要不断购置大量的存储设备来应对不断增长的存储需求。然而，单纯地提高存储容量，这似乎并不能从根本解决问题。首先，存储设备的采购预算越来越高，大多数企业难以承受如此巨大的开支。其次，随着数据中心的扩大，存储管理成本、占用空间、制冷能力、能耗等也都变得越来越严重，其中能耗尤为突出。再者，大量的异构物理存储资源大大增加了存储管理的复杂性，容易造成存储资源浪费和利用效率不高。因此，我们需要另辟蹊径来解决信息的急剧增长问题，堵住数据“井喷”。

高效存储理念正是为此而提出的，它旨在缓解存储系统的空间增长问题，缩减数据占用空间，简化存储管理，最大程度地利用已有资源，降低成本。目前业界公认的五项高效存储技术分别是数据压缩、重复数据删除、自动精简配置、自动分层存储和存储虚拟化。数据压缩和重复数据删除是实现数据缩减的两种关键技术。简而言之，数据压缩技术通过对数据重新编码来降低冗余度，而重复数据删除技术侧重于删除重复的文件或数据块，从而实现数据容量缩减的目的。

本文将介绍数据压缩和重复数据删除技术的效率对比。



### Lempel-Ziv系列压缩编码算法:

数据压缩的起源可以追溯到信息论之父香农（Shannon）在1947年提出的香农编码。1952年霍夫曼（Huffman）提出了第一种实用性的编码算法实现了数据压缩，该算法至今仍在广泛使用。1977年以色列数学家Jacob Ziv 和Abraham Lempel提出了一种全新的数据压缩编码方式，Lempel-Ziv系列算法（LZ77和LZ78，以及若干变种）凭借其简单高效等优越特性，最终成为目前主要数据压缩算法的基础。LZ系列算法属于无损数据压缩算法范畴，采用词曲编码技术实现，目前主要包括LZ77、LZSS、LZ78和LZW四种主流算法。

Lempel-Ziv系列算法的基本思路是用位置信息替代原始数据从而实现压缩，解压缩时则根据位置信息实现数据的还原，因此又被称作“字典式”编码。目前存储应用中压缩算法的工业标准（ANSI、QIC、IETF、FRF、TIA/EIA）是LZS（Lempel-Ziv-Stac），由Stac公司提出并获得专利，当前该专利权的所有者是Hifn, Inc.。

数据压缩的应用可以显著降低待处理和存储的数据量，一般情况下可实现2:1 ~ 3:1的压缩比。

### 文件级重复数据删除:

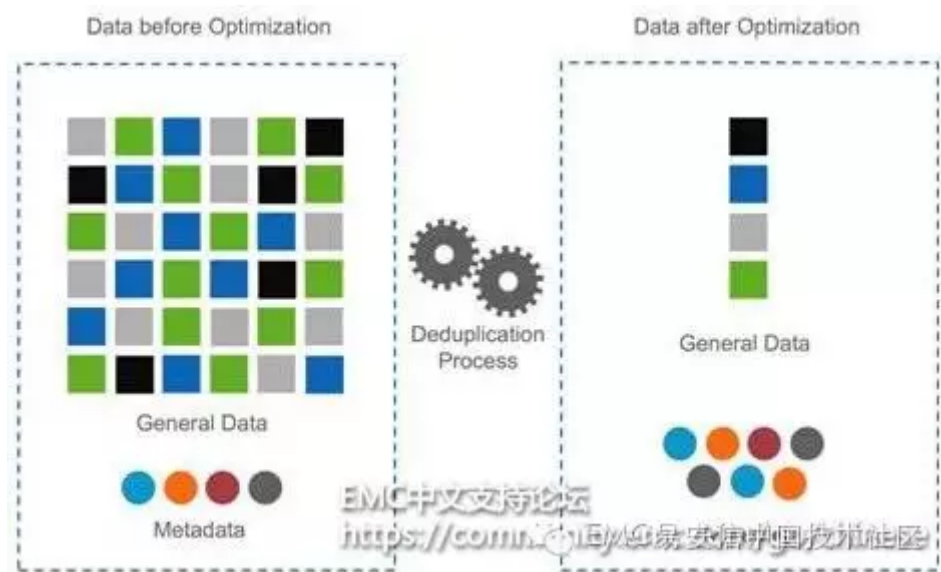
文件级消重，通常也称为单实例存储（Single Instance），原理很简单。在文件系统中检查并判断两个文件是否完全相同，如果发现两个相同的文件，其中一个就会被指向另一个文件的指针所取代。

文件集消重的数据缩减效率通常在3:1的压缩比。

### 数据块级重复数据删除：

在备份、归档等实际的存储实践中，人们发现有大量的重复数据块存在，既占用了传输带宽又消耗了相当多的存储资源：有些新文件只是在原有文件上作了部分改动，还有某些文件存在着多份拷贝，如果对所有相同的数据块都只保留一份实例，实际存储的数据量将大大减少--这就是重复数据删除技术的基础。这一做法最早由普林斯顿大学李凯教授（DataDomain的三位创始人之一）提出，称之为全局压缩（Global Compression），并作为容量优化存储推广到商业应用。

重复数据删除是一种数据缩减技术，可对存储容量进行有效优化。它通过删除数据集中重复的数据，只保留其中一份，从而消除冗余数据，其原理如下图所示。消重技术可以有效提高存储效率和利用率，数据可以缩减到原来的1/20 ~ 1/50。这种技术可以很大程度上减少对物理存储空间的需求，减少传输过程中的网络带宽，有效节约设备采购与维护成本。



数据块级的消重技术可以提供更高的数据消重率，因此目前主流的重复数据删除产品都是数据块级的。这些消重技术将文件分割成定长或变长的数据块，采用MD5/SHA1等Hash算法为数据块计算指纹（FP, Fingerprint）。可以同时使用两种及以上hash算法计算数据指纹，以获得非常小的数据碰撞发生概率。具有相同指纹的数据块即可认为是相同的数据块，存储系统中仅需要保留一份。这样，一个物理文件在存储系统就对应一个逻辑表示，由一组FP组成的元数据。当进行读取文件时，先读取逻辑文件，然后根据FP序列，从存储系统中取出相应数据块，还原物理文件副本。

### 数据压缩与重复数据删除对比分析：



数据压缩和重复数据删除技术都着眼于减少数据量，其差别在于数据压缩技术的前提是信息的数据表达存在冗余，以信息论研究作为基础；而重复数据删除的实现依赖数据块的重复出现，是一种实践性技术。然而，通过上面的分析我们发现，这两种技术在本质上却是相同的，即通过检索冗余数据并采用更短的指针来表示来实现缩减数据容量。它们的区别关键在于，消除冗余范围不同，发现冗余方法不同，冗余粒度不同，另外在具体实现方法有诸多不同。

### 1. 消除冗余范围

数据压缩通常作用于数据流，消除冗余范围受到滑动窗口或缓存窗口的限制。由于考虑性能因素，这个窗口通常是比较小的，只能对局部数据产生作用，对单个文件效果明显。重复数据删除技术先对所有数据进行分块，然后以数据块为单位在全局范围内进行冗余消除，因此对包含众多文件的全局存储系统，如文件系统，效果更加显著。如果把数据压缩应用于全局，或者把重复数据删除应用于单个文件，则数据缩减效果要大大折扣。

### 2. 发现冗余方法

数据压缩主要通过串匹配来检索相同数据块，主要采用字符串匹配算法及其变种，这是精确匹配。重复数据删除技术通过数据块的数据指纹来发现相同数据块，数据指纹采用hash函数计算获得，这是模糊匹配。精确匹配实现较为复杂，但精度高，对细粒度消除冗余更为有效；模糊匹配相对简单许多，对大粒度的数据块更加适合，精度方面不够。

### 3. 冗余粒度

数据压缩的冗余粒度会很小，可以到几个字节这样的小数据块，而且是自适应的，不需要事先指定一个粒度范围。重复数据删除则不同，数据块粒度比较大，通常从512到8K字节不等。数据分块也不是自适应的，对于定长数据块需要事先指定长度，变长数据分块则需要指定上下限范围。更小的数据块粒度会获得更大的数据消冗效果，但计算消耗也更大。

### 4. 性能瓶颈

数据压缩的关键性能瓶颈在于数据串匹配，滑动窗口或缓存窗口越大，这个计算量就会随之大量增加。重复数据删除的性能瓶颈在于数据分块与数据指纹计算，MD5/SHA-1等hash函数的计算复杂性都非常高，非常占用CPU资源。另外，数据指纹需要保存和检索，通常需要大量内存来构建hash表，如果内存有限则会对性能产生严重影响。

### 5. 数据安全

这里的数据压缩都是无损压缩，不会发生数据丢失现象，数据是安全的。重复数据删除的一个问题是，利用hash产生的数据块指纹可能会产生的碰撞，即两个不同的数据块生成了相同的数据指纹。这样，就会造成一个数据块丢失的情况发生，导致数据发生破坏。因此，重复数据删除技术存在数据安全隐患。

### 6. 应用角度

数据压缩直接对流式数据进行处理，不需要事先对全局信息进行分析统计，可以很好地利用流水线或管道方式与其他应用结合使用，或以带内方式透明地作用于存储系统或网络系统。重复数据删除则需要对数据进行分块处理，需要对指纹进行存储和检索，需要对原先物理文件进行逻辑表示。如果现有系统要应用这种技术，则需要对应用进行修改，难以做到透明实现。目前重复数据删除并不是一个通常功能，而更多地以产品形态出现，如存储系统、文件系统或应用系统。因此，数据压缩是一种标准功能，而重复数据删除现在还没有达到这种标准，应用角度来看，数据压缩更为简单。

