

Veritas Volume Replicator实现容灾

VERITAS Volume Replicator, 简称为VVR, 是VERITAS大型集成化容灾解决方案的一部分, 用来保证关键业务最大的可用性。作为一个与VERITAS的工业化标准产品Volume Manager紧密集成的模块, VVR将在任何基于裸设备和文件系统的数据环境中进行数据卷的复制。

设定Site 1为生产系统, Site 2为容灾系统。VERITAS Volume Manager (VxVM) 将在物理磁盘上建立多个或一个逻辑卷(Volume)。以裸设备的方式使用卷, 或在卷上建立文件系统。将数据(特别是需要进行远程复制的相关文件系统、数据库)存放在卷上。

VVR负责远程数据复制。VVR与VxVM完全集成在一起, 复制基于Volume进行。复制的数据可以是数据库中的数据(文件方式或裸设备方式)和文件。

将各个业务系统中需要进行远程复制的多个或一个卷定义为一个Replicated Volume Group(简称RVG); 在Site 1定义RLINK, 指向Site 2, 每个RVG定义一个RLINK。RLINK是任意的TCP/IP网络。Storage Replicator Log(简称SRL)是VVR中的重要部件。将数据复制各方的某个卷定义为一个SRL。需要复制的数据首先要写入SRL, 然后传到异地。VVR通过SRL保证数据复制严格按照写顺序进行。

在上述机制的作用下, VVR可完成任意距离的数据复制工作。

复制模式选择

数据容灾技术, 又称为异地数据复制技术, 按照其实现的技术方式来说, 主要可以分为同步传输方式和异步传输方式(各厂商在技术用语上可能有所不同)。下面, 我们将主要按同步传输方式和异步传输方式对数据容灾展开讨论。

同步数据复制

有关同步数据复制, 在传统意义上讲, 就是通过容灾软件, 一般是指异地数据复制软件, 将本地生产数据通过某种机制复制到异地。从广义上讲, 同步数据复制是指在异地建立起一套与本地数据实时同步的异地数据。

不同的异地数据复制技术的实现方式是不同的, 包括:

1. 基于主机逻辑卷层的同步数据复制方式;
2. 基于磁盘系统I/O控制器的同步数据复制方式;
3. 基于主机逻辑卷磁盘镜像功能的异地磁盘镜像方式;

异步数据复制

异步数据复制主要是针对“线路带宽和距离能保证完成数据复制过程, 同时, 希望异地数据复制不影响生产系统的性能”这样的要求下提出来的。考虑异步数据复制, 应该注意到以下几个技术条件和事实。

1. 带宽必须能保证将本地生产数据基本上完全复制到异地容灾端, 还要考虑距离对传输能力的影响。
2. 异地容灾端数据虽然落后, 但必须保证该数据库内在的数据完整性(一致性)、可用性, 否则, 这种数据复制就没有应用价值了。
3. 异地容灾端数据会比本地生产端数据落后一定时间, 这个时间随采用的技术, 带宽、距离、数据流特点的不同而不同。一般而言, 软件方式的数据复制技术具有完整的数据包的排队和断点重发机制, 在灾难情况下可以保证灾难时间点的数据一致性。
4. 异步容灾基本不影响本地系统性能。

与同步传输方式相比, 异步传输方式对带宽和距离的要求低很多, 它只要求在某个时间段内能将数据全部复制到异地即可, 同时异步传输方式也不会明显影响应用系统的性能。其缺点是在本地生产数据发生灾难时, 异地系统上的数据可能是以前的数据, 即最近时间内的交易会丢失。(注: 一个经过仔细计算和规划的系统, 才能保证其数据丢失最少。)

通过异步传输模式进行异地数据复制的技术, 包括:

- 1 基于主机逻辑卷的数据复制方式
- 1 基于磁盘系统I/O控制器的数据复制方式

基于主机逻辑卷 (Volume) 的数据复制方式

VERITAS VVR (Volume Replicator) 通过基于Volume和Log的复制技术, 保证在任何时刻本地系统发生自然灾害时, 在异地的数据仍是可用的。

VERITAS VVR在异步模式下采用了Log技术来跟踪未及时复制的数据块, 这个Log是一个先到先服务的堆栈, 每一笔I/O处理都会首先被放进这个Log, 并按到达先后顺序被复制到异地服务器系统。

整个I/O和复制的过程如下:

1. 本地主机系统发出第一个I/O请求A到逻辑卷;
2. 逻辑卷对本地磁盘系统发出I/O请求;
- 2'. 在往本地磁盘系统I/O的同时, 逻辑卷向本地磁盘系统上的VVR Log发出相同的写请求;
3. 本地磁盘系统完成I/O操作; 并通知逻辑卷“I/O完成”;
- 3'. VVR完成针对这个I/O的有关操作, 并通知逻辑卷;
4. 逻辑卷向主机确认“I/O完成”, 然后, 主机系统发出第二个I/O请求B。

服务器的另一个进程: VVR的进程, 负责将Log队列中的I/O复制到异地服务器。这个过程和上面的I/O过程在时间上无关。如上图中的标记: “I”和“II”。

I: 本地VVR进程从Log队列中取出最先到达的I/O, 复制到异地服务器

II: 异地服务器接收到本地服务器VVR发出的I/O请求, 将相应数据写到异地磁盘系统, 然后, 通知本地系统VVR进程, 要求下一个I/O。

这里, 跟踪未及时复制的数据块的Log技术是保证异地数据可用的必要条件。一个数据库的I/O是有严格顺序的, 这个顺序是保证数据库完整性的必要条件, 一个完整性被破坏的数据库一般是不可用的, 比如根本无法启动、打开该数据库, 且是无法修复的。本地数据库的完整性是由数据库本身来维护的。当一个数据库被实时复制到异地时, 要保证异地数据库的完整性, 必然保证在异地磁盘I/O上的I/O顺序和本地I/O顺序完全相同, 否则, 异地数据库的完整性就无法保证。

VERITAS VVR采用的I/O控制机制是支持先到先服务的Log技术, 因此, 不管异地数据比本地数据落后多少时间, 都能保证异地数据库数据的一致性。比如: 本地系统在12:00时发生自然灾害, 由于部分数据未被及时复制到异地, 如有1分钟的数据未完成复制, 那么在异地系统上存在11:59分钟以前的所有数据, 且这个数据库是可用的。