

## 1.1 备份的组成

能否安全有效的备份取决于以下三个因素：

- 存储备份数据的介质

存储硬件的质量与性能在整个备份过程中是至关重要的，它是能否进行高质量备份的关键所在。目前用于备份的存储主要有磁盘设备和磁带设备。磁盘设备具有快速的读写和快速搜索能力，适合于快速的小数据量备份（几百 G 亿下），但不适合于数据的较长期离线保存；而磁带设备具有大容量、每兆字节价格低和便于离线存放的特征，适合于大数据量的备份（1T 以上），但是执行备份和恢复的速度相对较慢。

介质是数据的负载物，可以是磁带也可以是硬盘。它的质量一定要有保证，使用质量不过关的介质无疑是拿自己的数据冒险。由于磁带是一种易耗品，有一定的使用寿命，使用者应当知晓何时需要更换新磁带。有些好的磁带会自带一个使用记录，当磁带使用次数超过了最高限制，会自动报警，提醒用户更换磁带。

- 备份的存储架构

目前外部存储的实现方式主要有三种，分别是：DAS 存储系统、NAS 存储系统、SAN 存储网络。

- 控制备份的软件

操作系统通常包括一些基本功能，但是为了达到更好的备份效果，最好使用专门的备份软件。优秀备份软件包括加速备份、防病毒、自动操作、灾难恢复等特殊功能，对于安全有效的数据备份是非常重要的。理想的备份软件与自动加载磁带机配合，可提供自动备份、自动磁带轮换、自动诊断和数据统计功能。

## 1.2 备份介质的选择：磁盘 OR 磁带

人们一提起数据备份，头脑中联想到的等价词就是“磁带”，可以说，磁带伴随着备份技术发展的每一个阶段，大家都即定俗成地将“数据备份”与“磁带”划上了等号。但是，用磁盘备份具有速度快、备份窗口小的优势，磁带不再是备份数据的唯一介质选择。

业界对磁带和磁盘的争论可算争论已久，自 ATA 磁盘在存储中的应用，成为了存储业界炙手可热的话题。一派人认为磁带将会被淘汰，而磁盘将会成为主要的备份介质。虽然有很多论据支持将磁带从磁带备份中换掉，由磁盘备份范例来替代，但磁带仍然是备份与归档系统中占据着不可缺少的一部分。

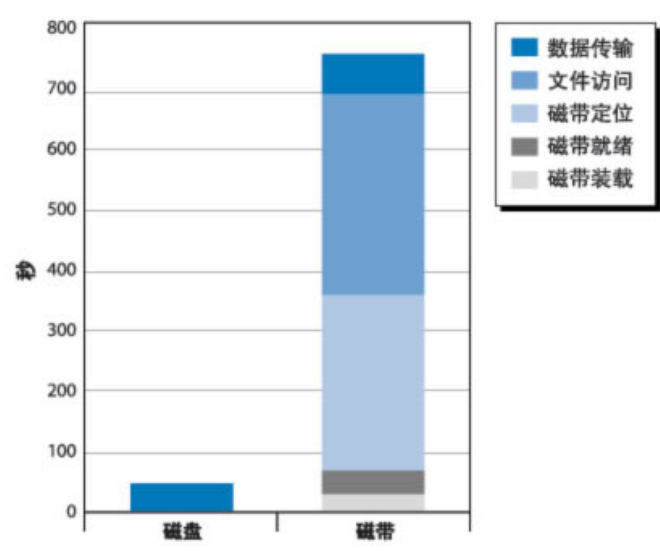
### 1.2.1 磁盘的优势

**恢复的速度更快：**磁盘驱动器的恢复的速度要比磁带驱动器快 - 两者之间是几秒或几分钟与几小时的差异。此外，磁盘支持随机访问和顺序访问，而磁带仅支持顺序访问。支持两种

类型可以更快地访问磁盘上的数据文件，提高了整体性能。

从多个磁带盒中恢复数据需要很多耗时的步骤。首先，磁带库必须安装每个磁带；每个磁带耗费的时间达到 1 分钟。然后必须要加载磁带，这还会占用 30 秒钟到几分钟。磁带必须定位到所需要的数据；平均访问时间是几分钟。其次，磁带必须重绕和卸载，这将会需要 30 秒钟到几分钟。在加载磁带之后，这个周期又再次重复。访问第一个字节所需的时间（TTFB）对于磁带来说是几秒钟到几分钟，但对于磁盘来说只有几毫秒。

**介质可靠性和数据可用性：** 磁盘系统的 RAID 保护功能增强了数据的可用性，防止磁盘驱动器出现故障时导致数据丢失，而磁带所引起的介质错误是很常见的事情，磁带机、磁头易出机械故障。维护磁带库中的磁带需要经过培训的人员。通过使用磁盘存储方式，IT 部门可以减少或消除对磁带的处理。



**效率高：** 因为 RAID 保护使磁盘本身更为可靠，所以在使用磁盘是要进行完整备份的次数更少，这就节省了网络和 CPU 的负担。磁带技术和磁带数据格式往往每几年就发生变化，这就迫使 IT 部门转换为新的介质。磁盘技术不存在这种转换，因为数据格式不会变化。虽然多路技术允许更快速地写入到磁带或磁盘，但从磁带上恢复多路化的数据要比从磁盘上恢复慢得多。此外，与同等容量的磁带库相比，新的容量更大的磁盘驱动器占用较少的地面空间。

**成本低：** 磁盘技术发展很快，成本不断下降，出现了各种低端的磁盘阵列，如 ATA 磁盘阵列。此外磁盘备份可降低维护磁带所需的管理成本。

磁带、磁盘备份比较列表		
备份介质	磁带	磁盘
物理特性	线性记录介质，无论读写数据都需要顺序操作，涉及过多机械物理操作	随机记录介质，读写数据都可以随即进行，没有过多机械操作
可靠性	没有校验操作，可靠性较差，一般采用多次备份的方式来提高可靠性	可靠性较高，先进方案还通过软件方式提供高级别校验，但是不可避免存在“带电介质”本身固有的可靠性隐患
远程传输	需要通过卡车运输等传统方式进行，效率较低	可以通过快速的网络传输方式进行，效率高，但是费用同样较高
总体价格	当备份数据量非常大时，价格优势十分明显	备份数据时间短、数据量不大时价格较低，高压缩比的方案成本优势更加明显

### 1.2.2 磁带的价值

归档。用户的备份数据分为两类，一类是需要随时可用的“热数据”，或者说在线备份数据；另外一类是保存起来以备不时之需的“冷数据”，或者说历史数据。

“热数据”能随时用于恢复和其它用途，同时热数据保存的时间短，数据量小；通常企业希望保存三个月的在线备份数据，数据首先备份到磁盘上，然后归档到磁带上，磁盘上备份企业的“热数据”，可以随时进行快速恢复和数据共享；

### 1.2.3 “冷数据”主要用于保存和归档历史数据，提供给用户进行特殊

的历史查询或者历史备份版本的恢复。冷数据保存时间长，数据量

大，但使用的机会少，企业通常需要保存 3 个月以上、甚至超过 1

年的历史数据。使用磁带来保存“冷数据”，充分利用了它的价廉，

容量大，可移动的优点。归档过程不对生产环境产生影响，也避免

了其备份恢复速度慢和可在备份 Oracle 数据库方面的特殊优势

DSG SnapAssure 不同于普通的备份软件如 Bakbone、Veritas、Legato 等对 Oracle 数据库的备份方式，SnapAssure 不使用 RMAN 作为 Oracle 数据库的备份接口，而是基于磁盘技术实现的“增强备份”系统，这是一种基于数据块级增量的物理备份技术。

SnapAssure 在被保护的 Oracle 数据库的系统上运行一个代理程序(AGENTD)，协调和需要备份系统的关系，获得需要备份的系统文件或者数据库文件所有的数据块发生的改变。对于不同的备份方式，如全备份、增量备份等，实现的机制与 RMAN 完全不同。

在对 Oracle 数据库做全备份时，SnapAssure 首先执行的是一个物理全备份，通过日志保证物理全备份结束后，所有数据文件、控制文件等的一致性，保证全备份版本是一个可以将数据库恢复且数据库可以正常启动的版本；

在增量备份时，当用户通过 SnapAssure 控制台下达备份指令后，运行在 Oracle 数据库所在系统的代理（或称客户端）直接将(至上次备份以来)发生改变的数据块备份到管理平台。在代理程序读取数据块的同时，将数据块压缩，通过网络传输到存储系统保存。所以一体化存储管理平台实际上只有两种备份类型：完全备份和增量备份。

每一次备份结束 SnapAssure 通过专有的文件系统形成逻辑上完整的备份数据版本。无论是完全备份还是增量备份，在 SnapAssure 上都形成逻辑上完整的备份版本。系统通过多种技术手段来保障备份数据的可靠性和一致性。

在需要进行数据恢复时，可以从任何一个需要恢复的版本进行数据恢复。在用户发送恢复指令后，SnapAssure 从备份数据集中提取所有该版本的数据文件，恢复到用户指定的系

统。同时和数据文件保存的日志文件、控制文件等也一起恢复。用户使用保存的数据文件、日志文件、归档日志文件，可以将数据库恢复到需要的时间点。

在存储设备方面，DSG SnapAssure 可以支持几乎所有的磁盘阵列设备。

SnapAssure 备份技术提供了灵活的升级能力、可管理性和可用性。提供数据的在线保护和粒状恢复，是全球存储界的优秀磁盘备份技术。

#### 1.2.4 SnapAssure 的特点

迪思杰的 SnapAssure 备份技术采用磁盘阵列作为备份主要介质，同时支持备份数据向带库转移的功能。SnapAssure 是为跨异构平台的服务器和客户端而设计的、企业级的、提供了分布式网络体系结构下数据存储管理解决方案。它通过网络中选定一台机器作为数据管理的备份服务器，在其它机器上安装客户端软件，从而将整个网络的数据自动地备份到与备份服务器相连的存储设备上，并在备份服务器上为各个备份客户端建立相应的备份数据的索引表信息，用来实现数据的全自动恢复操作提示。SnapAssure 备份技术提供了灵活的升级能力、可管理性和可用性。提供数据的在线保护和粒状恢复，是全球存储界的优秀磁盘备份技术

在满足用户的新一代备份系统的基本建设目标方面，有以下杰出的特点：

- 高速备份和恢复性能：迪思杰的备份技术，可以提高传统备份和恢复速度数倍（一般约 3-8 倍之间）。
- 备份数据的高可靠性：首先磁盘具有 raid 机制保护、错误报告等机制，所以用来作为一级的备份设备，能够比带库提供更大的可靠性保证。同时，迪思杰独创的备份数据监测技术，可以高速地实现备份数据的正确性的检查，避免系统在恢复时发现数据不正确的局面。
- 低系统干扰：独创数据流压缩技术，独创的数据库数据、日志分析技术，大幅度减少备份性能受带宽资源限制，大幅度降低对生产系统的干扰，大幅度降低磁盘存储容量的需求；
- 简单备份策略、完全自动管理：迪思杰首推智能完全备份技术，通过采用独创压缩版本文件系统，智能管理备份的数据。这项技术，使得日常的备份任务变得非常简单，可以将全备份、增量备份简单到只是通过增量数据备份就能获得完整的备份版本；
- 支持 oracle 在归档和非归档模式下的备份：支持 ORACLE 数据库在非归档日志模式（No archive log）下的在线热备份，也支持归档备份。

#### 1.2.5 与 VTL 的差别

DSG SnapAssure 推荐用户采用磁盘阵列作为备份介质，为了节省成本可选择采用 SATA 盘阵列，但业界也推出了虚拟磁带库（VTL）技术用于磁盘备份。

虽然二者方案都采用磁盘阵列作为基础介质，但有很大的差别：

- 虚拟带库技术是在磁盘阵列上增加软件将其模拟成为磁带驱动器的接口标准，以期利用磁盘阵列 I/O 快的特性来提高备份和恢复的效率。
- DSG SnapAssure 之所以采用磁盘阵列是为了利用磁盘阵列的随即读取功能，而不是简单因为其速度快，从而从备份的实现原理上做了根本性的改变。

从以下几个角度，我们可以分析二者的差别：

	VTL 技术+传统备份软件	DSG SnapAssure+磁盘阵列
增量备份的效率	<p>■ <b>有所提高，但效率提高幅度不大。</b></p> <p>因为增量备份慢的原因主要不取决于存储介质是带库还是磁盘，主要原因是 RMAN 的增量备份原理造成的。RMAN 做增量备份时需要对数据块进行扫描，即使只有几个块发生了变化，<b>RMAN 也需要花费很长时间进行全库扫描。</b></p> <p>所以即使采用 VTL，现有的备份软件依然无法提高备份效率，增量备份效率依然很低。</p>	<p>■ <b>大幅度提升增量备份效率：1TB 的数据库增量备份时间可控制在 30 分钟左右。</b></p> <p>备份效率之所以提高，不是因为用磁盘阵列做存储介质，而是因为 DSG SnapAssure 采用了无需扫描的增量备份技术。</p> <p>在增量备份时，SnapAssure 不需要从扫描全库的数据中去发现变化的块。</p>
备份策略	<p>■ <b>不能从根本上得到改善：</b></p> <p>因为现有备份软件是因为增量备份性能低下，所以每天才只备份归档日志，而不是按照宣称的那样每天都做增量备份或者差量备份。</p> <p>所以即使改成了 VTL，那么备份策略依然不能改变。</p>	<p>■ <b>根本上改善：每天都会进行一次增量备份（SnapAssure 叫合成全备份）。</b></p> <p>因为 DSG SnapAssure 增量备份的速度提高了，每天的增量备份时间在 1 小时左右。所以，SnapAssure 每天都将进行数据文件和归档日志的双备份。</p>
恢复效率	<p>■ <b>Restore 过程将有所提高，但 Recover 过程无法提升：</b></p> <p>采用 VTL 后，Restore 一个文件或所有文件的效率相对于磁带库而言将会有很大提升。</p> <p>但是对于 Recover 过程却无法提升，因为如果只备归档日志，那么在数据恢复时需要 Recover 的日志量就非常大，例如在电信行业可能达到 600~700GB 左右，这么大的日志量 Recover 的时间也是需要非常长时间的。</p>	<p>■ <b>Restore 和 Recover 都提升：</b></p> <p>DSG SnapAssure 采用合成备份概念，每次备份都是一个全备，无论是恢复周日的数据，还是恢复周六的数据，都只需要从一次备份版本中恢复。</p> <p>同时，在 Recover 过程中，由于 SnapAssure 每天都有一个合成全备，那么在 Recover 过程中，最多只需要恢复 1 天的日志，以电信行业为例，&lt;100GB。</p>
备份数据的可验证性	<p>■ <b>几乎没有。</b></p> <p>VTL 技术备份的数据如果要确认是否能够恢复，最彻底的办法就是将数据恢复并 Open 一次。</p> <p>这一点与现有的带库备份没有什么差别</p>	<p>■ <b>支持。</b></p> <p>DSG SnapAssure 提供了备份数据可打开验证的手段，从备份<b>服务器</b>上可直接打开备份的数据，让 Oracle 数据库来验证备份数据是否可用，是否可以恢复。</p>

表恢复	<p>■ <b>不支持。</b></p> <p>即使采用 VTL 技术，也只能支持物理恢复，比如全恢复、表空间恢复、数据文件恢复等。</p> <p>但对于因为误操作造成的 drop table, truncate table, drop table, truncate table 以及其他的逻辑误操作造成的损坏，VTL 的恢复过程将非常复杂。</p>	<p>■ <b>支持。</b></p> <p>DSG SnapAssure 提供单表直接恢复功能，能从备份系统上直接恢复某张表。</p> <p>这是其他备份软件都不具备的功能。</p>
非归档备份	<p>■ <b>不支持。</b></p>	<p>■ <b>唯一支持。</b></p>