**DRBD安装配置、工作原理及故障恢复**

一、DRBD简介

DRBD的全称为：Distributed ReplicatedBlock Device(DRBD)分布式块设备复制,DRBD是由内核模块和相关脚本而构成，用以构建高可用性的集群。其实现方式是通过网络来镜像整个设备。你可以把它看作是一种网络RAID。它允许用户在远程机器上建立一个本地块设备的实时镜像。

二、DRBD是如何工作的呢?

(DRBD Primary)负责接收数据，把数据写到本地磁盘并发送给另一台主机(DRBD Secondary)。另一个主机再将数据存到自己的磁盘中。目前，DRBD每次只允许对一个节点进行读写访问，但这对于通常的故障切换高可用集群来说已经足够用了。有可能以后的版本支持两个节点进行读写存取。

三、DRBD与HA的关系

一个DRBD系统由两个节点构成，与HA集群类似，也有主节点和备用节点之分，在带有主要设备的节点上，应用程序和操作系统可以运行和访问DRBD设备（/dev/drbd\*）。在主节点写入的数据通过DRBD设备存储到主节点的磁盘设备中，同时，这个数据也会自动发送到备用节点对应的DRBD设备，最终写入备用节点的磁盘设备上，在备用节点上，DRBD只是将数据从DRBD设备写入到备用节点的磁盘中。现在大部分的高可用性集群都会使用共享存储，而DRBD也可以作为一个共享存储设备，使用DRBD不需要太多的硬件的投资。因为它在TCP/IP网络中运行，所以，利用DRBD作为共享存储设备，要节约很多成本，因为价格要比专用的存储网络便宜很多；其性能与稳定性方面也不错

四、DRBD复制模式

协议A：

异步复制协议。一旦本地磁盘写入已经完成，数据包已在发送队列中，则写被认为是完成的。在一个节点发生故障时，可能发生数据丢失，因为被写入到远程节点上的数据可能仍在发送队列。尽管，在故障转移节点上的数据是一致的，但没有及时更新。这通常是用于地理上分开的节点

协议B：

内存同步（半同步）复制协议。一旦本地磁盘写入已完成且复制数据包达到了对等节点则认为写在主节点上被认为是完成的。数据丢失可能发生在参加的两个节点同时故障的情况下，因为在传输中的数据可能不会被提交到磁盘

协议C：

同步复制协议。只有在本地和远程节点的磁盘已经确认了写操作完成，写才被认为完成。没有任何数据丢失，所以这是一个群集节点的流行模式，但I / O吞吐量依赖于网络带宽

一般使用协议C，但选择C协议将影响流量，从而影响网络时延。为了数据可靠性，我们在生产环境使用时须慎重选项使用哪一种协议

四、 DRBD工作原理图

DRBD是linux的内核的存储层中的一个分布式存储系统，可用使用DRBD在两台Linux服务器之间共享块设备，共享文件系统和数据。类似于一个网络RAID-1的功能，如图所示：

