

Replication is a widely used Disaster Recovery strategy that allows for the secondary node to failover and start fielding inline I/Os from the application immediately whenever there is a disaster in the primary node. In general, failover operations should only be performed in cases where the host (application server) is unable to reach the primary node.

Tech Sheet
08 15 2007

Overview of Failover Operations in StorTrends® iTX

In StorTrends iTX, failover is a quick-recovery operation for minimum downtime (small Recovery Time Objective, or RTO) in response to a disaster. The command to initiate this procedure must be executed in the original secondary storage appliance's ManageTrends GUI, or through CLI commands. Once the failover command is received, the original secondary node converts itself into the new active primary. After this point, the new primary will also initiate communication with the original primary (new secondary) node, and begin to update the original primary with any changes in the replication state.

During recovery and the subsequent failover, the original primary appliance may not be accessible from the secondary node. It is often possible that a link failure has occurred between the host and the primary node, and the administrator opted to perform a failover in order to reduce the down-time for the application. However, the secondary node may still be able to connect to the primary node through

a different network, and as such the original primary node will modify its status to become the "new" secondary node immediately, and replication will then proceed in the reverse direction. In the case of asynchronous replication a "failback join" operation must be performed on the active primary before replication can resume.

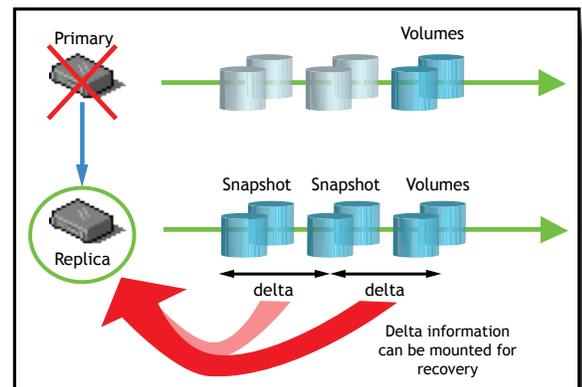
In cases where the original primary node remains down due to some catastrophic failure and hence is unreachable from the new primary node, the new primary will continue attempting to reach the original primary until it returns, and will log replication writes in "tabbing mode" until then. Asynchronous replication uses snapshots for the logging of data changes in a similar way.

Once the original primary is recommissioned, the new primary will perform a **handshake** with it. In order to perform this handshake and resolve "split-brain" conflicts where both nodes declare themselves as primary, the

concept of an **epoch number** is employed. In short, when replication is first started, the epoch number of replication is set to 0, and whenever a failover or a failback is performed, the epoch is incremented by 1.

Thus when the original primary comes back from an unplanned outage, the new primary (which has a higher epoch number) will be considered the correct active primary node, and the original primary will demote itself to secondary, and also update its own epoch number.

After this handshake is performed, the



new primary node will resynchronize the (original primary) new secondary node

Tech Sheet

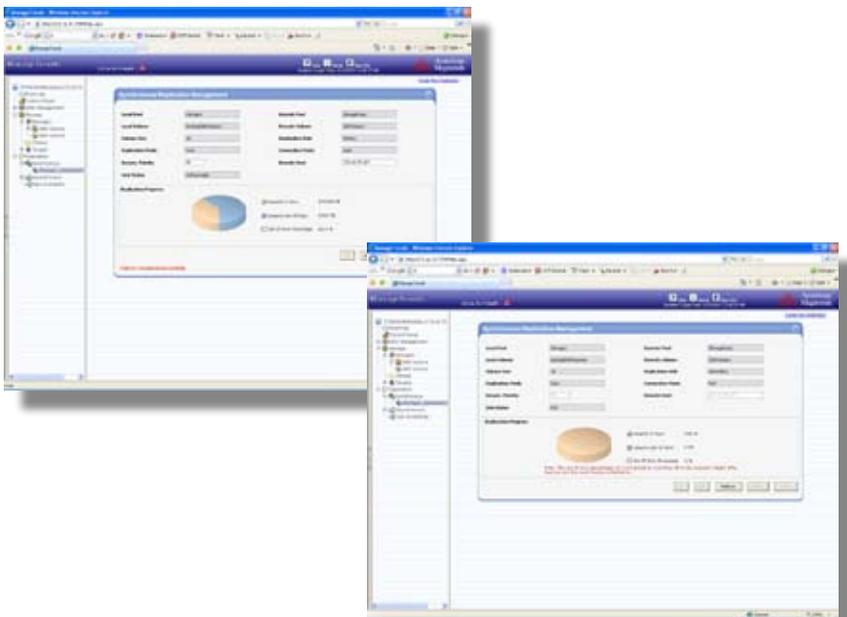
with all the tabbed I/Os that were logged during the time it was down. Until the original primary is fully synchronized with the active primary node, it cannot be failed-back to (i.e., restored to its original role as primary).

While generally unadvisable, as mentioned earlier, it is worth mentioning that it is nonetheless theoretically possible to failover when both nodes are up and the primary is reachable from the host. However, this operation does not add much value, and in cases where this must be done, it is recommended to do a clean logout from any running applications before performing a failover.

Failover Operation Procedure in StorTrends iTX

Initiating failover operations is a simple process. The user simply clicks on the “failover” button in the ManageTrends control console of the original secondary node, and failover is automatically handled by StorTrends iTX from this point forward. In the case of synchronous replication, once the failover is complete, the original secondary takes over as the new primary in replication operations. For asynchronous replication, this will occur after the fallback join is performed.

Screenshots of the ManageTrends® web-based GUI for failover operations are shown here:



ManageTrends: Screenshots of Failover Complete and Original Primary Resynchronization Complete

A sample CLI command for the failover operation of a synchronous replication pair is also shown here:

```
[root@ITX0030488422ee ~]# cli perform -t failover -l Storage1_SANVolume1 syncreplica
```

Why AMI?

AMI offers a wide array of disaster recovery and high availability products for your business needs. AMI and its partners provide services that range from storage needs analysis to the design and implementation of a custom disaster recovery solution. We can help your business plan for when things are at their worst while reducing the cost and complexity of your storage environment. For more information on AMI StorTrends solutions, visit www.StorTrends.com, email to sales@ami.com, or call (800) U.Buy.AMI.



American Megatrends Inc. | www.ami.com
6145-F Northbelt Parkway
Norcross GA 30071 | t: 770.246.8600
Sales & Product Information
sales@ami.com | t: 800.828.9264
Technical Support
support@ami.com | t: 770.246.8645