

High Availability Data Storage and Instant DR

Modern business's need high availability access to their data. When data becomes unavailable the business usually stops operating - and if the data is actually lost the cost and damage to the business may be even terminal; such is the continually increasing reliance on IT systems to run businesses.

The business's application servers can be clustered or virtualised using products such as VMware's ESX Server or Microsoft's Virtual Server to optimise the use of server processing power and protect from individual server hardware failures. The application can run on a virtual server which is a pool of physical server machines. If a physical machine fails then another can replace it quickly and that critical business application can be back up and running in a matter of minutes.

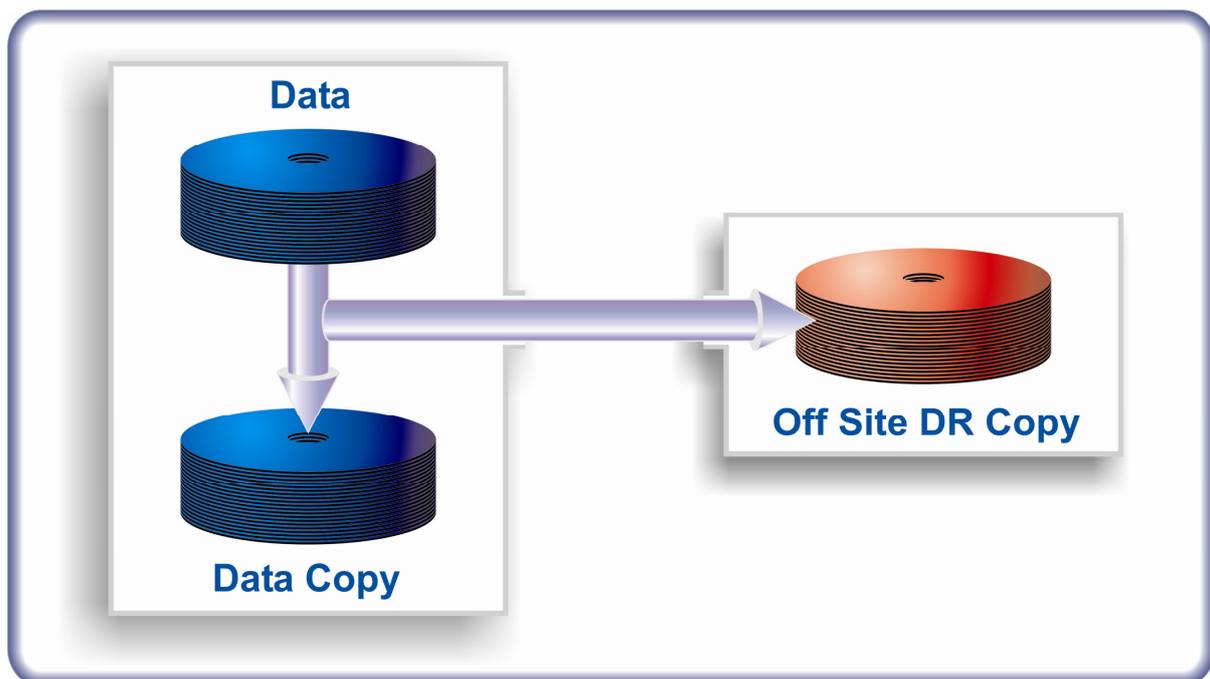
At the same time data also needs to be continually available for those servers (or virtual servers). Data needs to be stored on a highly reliable platform, be easy to back up and restore, and work with the type of applications, databases, and operating systems that the business uses.

There are two clear needs when it comes to business data:

1. It must always be available.
2. If there is a real problem then data needs to be recoverable quickly and simply to enable the business to get back up and running.

High Availability Data Solution

One increasingly popular and now more cost effective solution is to have highly available resilient storage replicated in real time at the primary site, and to have a second backup (Disaster Recovery) copy located off site somewhere.



Often the link to the off site copy has a restricted bandwidth so it is important that any data replication to the off site location makes the absolute best use of the available bandwidth and ensures that the off site data is as up to date as possible.

The ideal solution is provided by using block based replication technology, where only new data block writes are replicated to the secondary unit. Block based storage technology is completely operating system independent and data independent so will operate with any file system, database applications such as SQL or Exchange.

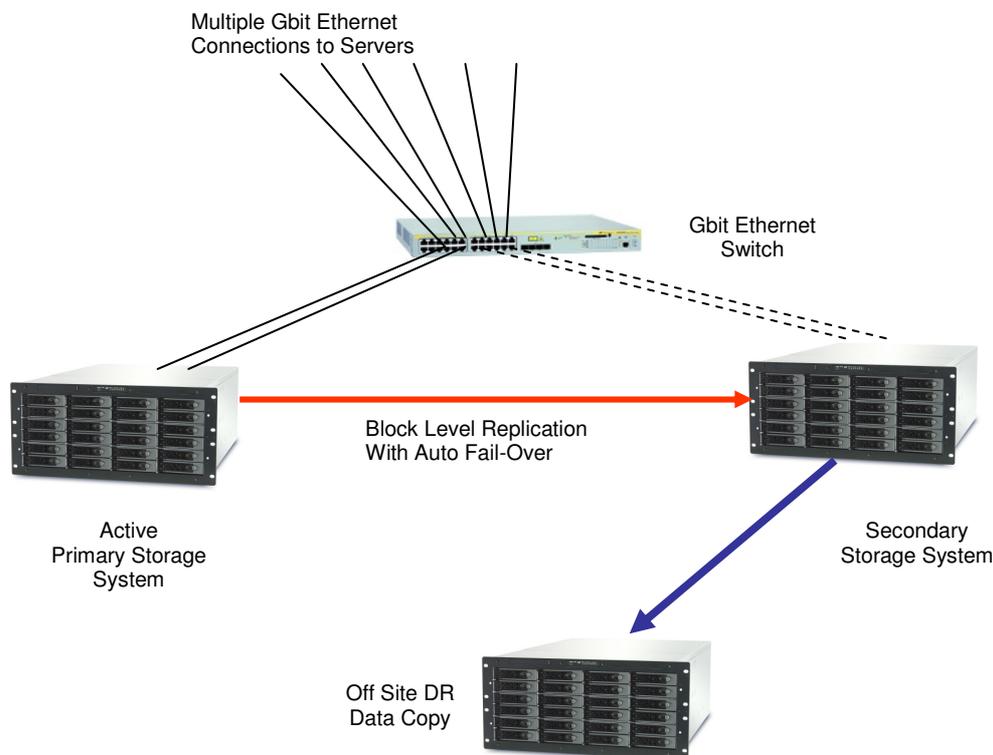
Block based means that only new writes of block data need be replicated over the link to the remote site, unlike file based replicated systems where even a small update to the file means that the entire file must be rewritten.

Data compression techniques can also be employed for block writes to the off site copy where the available bandwidth is limited (below 10Mbps typically).

Tertiary Replication System

Here is an example of a real data storage system providing very high data availability on the main site and an off site copy located at a remote site for DR purposes, all in real time at the block level using high performance iSCSI storage technology.

Two units at the main site provide an auto fail-over pair to ensure business continuity even in the unlikely event of a major failure with the primary storage system.



iSCSI storage technology is ideal for delivering such a system since it utilises low cost Gbit Ethernet. Multiple Gbit links or very high performance 10Gbit Ethernet may be used to provide the data bandwidth required for your application and to allow easy expansion as your business data capacity needs grow.

Coupled with automated snapshot features to enable quick and simple recovery of accidentally deleted files or to recover from a virus attack, asynchronous block level replication provides a near bullet proof protection system for data storage.

Automatic Failover

In the system shown above, under normal operating conditions data reads and writes are made to the primary active storage system. All new data block writes (only) are also replicated to the secondary passive storage system usually located somewhere else on the same site. Should a major problem develop with the active primary unit, the secondary automatically takes over to ensure continuous access to data and this is done so quickly that servers that have been provisioned with storage from the primary system will not even be aware that there has been a problem. This is achieved by the secondary unit taking over the same IP address as the primary unit in the event of a failure – no intervention at the server level is required and the whole process is completely seamless and takes a matter of seconds.

With a high bandwidth link between primary and secondary on the same site, a real time replicated copy of the data is maintained at all times.

At the same time, the secondary is also replicating new block writes to a third off site unit. Here the link to the off site unit will undoubtedly be slower than the main network and so a temporary buffer is used to allow new writes to be queued and written as quickly as possible to the remotely located unit, so making the best use of the bandwidth available.

If the worst should happen and the main site is lost completely, then the 3rd off site storage system provides a near real time copy of the main sites data, subject only to the limitations of the bandwidth from the main site.

Cost Effective Storage Capacity

With disk storage capacity ever increasing, and with the cost per Terra Byte falling all the time, such replicated systems are now providing a cost effective way of providing ultimate fully replicated protection for data which would previously been a very expensive proposition. Enterprise class disk drives with unformatted capacities of 1TByte designed for 7x24 hour continuous use in RAID systems are now shipping. Capacity has never been so affordable. The same disk technology boasts MTBF's of over 1 million hours, hard to believe for what is still a mechanical device.

Even with replication, it is still highly desirable for the storage systems themselves to be extremely reliable and use the usual hot swap technology to allow redundancy in the disks, power supplies and cooling parts without needing the backup of failover to protect from data loss.

RAID Data Protection

Certainly the advanced data guard protection offered by RAID 6 should be a prerequisite when choosing the right data storage system. RAID 6 protects the data even in the (very) unlikely event of two disks failing. This has become more important since as storage arrays become larger, the rebuild time to replace a failed drive may be quite lengthy and in more conventional systems running in RAID 5, a second disk

failure, however unlikely during the rebuild time, would result in data loss in a non replicated system – RAID 6 completely alleviates this while still providing the high read / write performance advantages of RAID 5.

RAID 6 data protection makes a significant contribution to providing a true high availability data system.

Performance

Performance-wise an iSCSI Storage Area Network is more than capable of providing and even exceeding the data throughput of traditional and more expensive fibre channel SANs. Fibre channel SANs running at 2Gbits/second and more recently 4Gbits/second are easily surpassed by multiple lower cost Gbit Ethernet links, with the added advantage of extra resilience provided by multi-path links. A single Gbit Ethernet link is capable of providing up to about 120Mbytes of data throughput. Multiple links can easily be aggregated to provide more bandwidth and failure of a single link is then almost insignificant.

For very intensive data storage applications 10Gbit Ethernet may be deployed for the link between the iSCSI storage system and the Ethernet switch, and then multiple Gbit links are used to connect to the servers. A single 10Gbit Ethernet link is capable of providing data bandwidth in excess of 1000Mbytes per second, far more than a fibre channel link, although this sort of bandwidth would only be required in the most demanding applications with a large number of performance servers being provisioned by the iSCSI storage system.

For more information on iSCSI technology and how to build your SAN using iSCSI technology please visit our website at www.westekuk.com