Infortrend

Using Multipathing Technology to Achieve a High Availability Solution

Table of Contents

Introduction

Greater storage system availability and improved performance are high priorities for business IT managers. This paper discusses the benefits of current multipathing solutions and how the technology increases end-to-end redundancy and achieves storage availability and high performance when used with a RAID array.

To meet the need for storage system availability, RAID technology helps by protecting the disk drives. However, if there is only a single path from the host to the storage device and one of the components in that path fails, the data will be unavailable.

Multipathing is a fault tolerant technique where there is more than one physical path between the CPU in a computer system and its mass storage devices. Redundant physical path components include buses, controllers, switches, and the bridge devices connecting them.

A simple example would be a SCSI disk connected to two SCSI controllers on the same computer or a disk connected to two Fibre Channel ports. If one controller, port or switch fails, the operating system can route I/Os through the remaining controller transparent to the application, except perhaps for incremental latency.

In addition to fault tolerance, a multipathing solution redistributes the I/O request load among the multiple paths between the host and the storage devices. This helps remove bottlenecks and balances workloads.

Multipathing Technology

The two major functions for a multipathing solution include fail-over protection and load balancing. With these two features, the system can achieve high availability and improved performance.

Fail-Over

A single connection between storage devices and hosts is shown in ***Figure 1***.

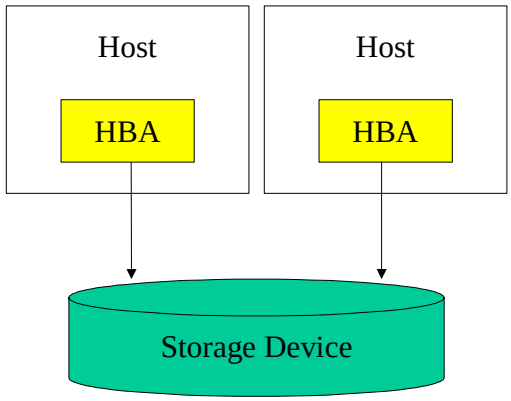

Figure 1: Single Path Connections between Devices

Any connected device can be associated with one host controller. There can be several host controllers and each controller can be responsible for several devices, but the data travels on one path only. This is the method used with IDE or parallel SCSI systems.

The data on the storage device can be protected by a RAID function but that doesn't increase the reliability of transmission between the storage devices and host. If the HBA or cable/bus fails, the connected devices can't be accessed.

Multipathing guards against potential damage resulting from a single connection failure. In ***Figure 2***, there is more than one connection between the host and storage devices. If any connection fails, the host can access the storage device via another path and maintain operation continuity.

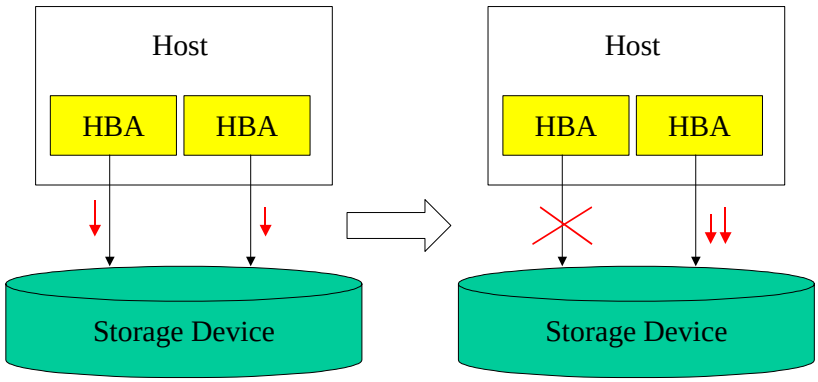

Figure 2: Fail-Over using Multipathing Devices

Load Balancing

In addition to providing fail-over protection, multipathing provides load balancing to increase system I/O performance. Load balancing is the redistribution of read/write requests in order to maximize throughput between the host and storage devices. If the operating system is aware of multiple data paths, it can use them to increase the throughput to or from the storage devices. This is especially important in high workload settings or other situations where consistent service levels are critical.

For example, the system shown on the left in ***Figure 3*** is not utilizing load balancing. The left path is saturated and is a bleneck for some applications running on the host.

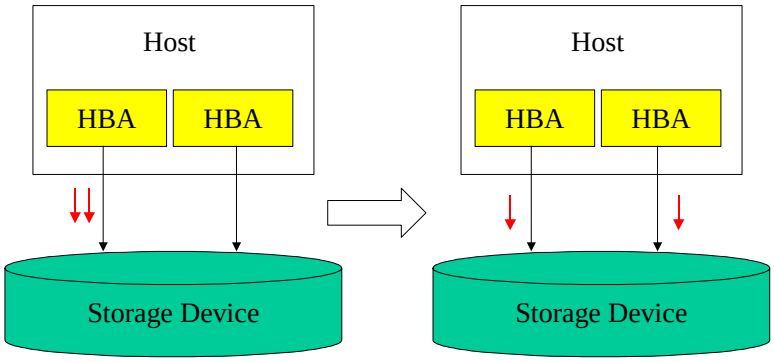

Figure 3: Load Balancing using Multipathing Devices

With load balancing capability, the host can distribute I/O traffic among all of the data paths. As shown on the system on the right in ***Figure 3***, some of the I/O traffic from the left path has been rerouted to the right path with extra available bandwidth. The bottleneck is removed and the total data traffic from the server is distributed equally.

Multipathing I/O Implementations
The three major types of providers for current multipathing solutions include storage subsystem providers, storage software providers, and OS platform providers. All three solutions operate on the same storage hardware.

In general, storage hardware providers offer complex multipathing storage management solutions on the host with device drivers that are optimized for their own subsystems and cannot be used with storage hardware from other manufacturers.

On the other hand, multipathing solutions implemented by a storage software provider or by the OS platform generally have simple path fail-over or may have more complex functionality in the volume management layer, device layer, or both. Customers can ensure better interoperability between multi-vendor storage devices by using a native, host-based multipathing solution; however, these host-based solutions do not provide the integrated cooperation capabilities for a specific storage device as provided by the storage device vendor.

Another approach is to provide a joint solution from both the host OS software and the storage device vendor. For example, the Microsoft MPIO solutions work in conjunction with device specific modules (DSMs) written by storage device vendors. This allows vendors to design specific hardware solutions that are tightly integrated with the OS platform so that customers can have both complex functionality and specific storage device capabilities for the best multipathing solution possible.

Storage Redundancy
Availability is achieved by using a multipathing function from the host side, while the RAID disk array provides the redundancy for fault tolerance on the storage side. There are two major redundancy modes for a disk array: active/active and active/passive.

Active/Active Mode
The "Active/Active" mode for disk arrays transfers I/Os to drives through multiple paths simultaneously, without causing any performance impact. Load balancing increases I/O throughput by utilizing the full bandwidth of all available paths.

Active/Passive Mode
"Active/Passive" is a mode for disk arrays in which controllers have their own Logical Unit Numbers (LUNs)/disks. The controller will only issue I/Os to its assigned LUNs. The controller is called the "primary" path and the alternate controller is called the "secondary" path to the LUN.

A disk array doesn’t allow access of a disk/LUN through multiple available paths when it is in active/passive mode. The disk array has to change ownership of the LUN between these two controllers if the LUN needs to be accessed via multiple paths automatically. This can cause immense performance degradation.

Since the load balancing software or drivers only know the status of current jobs on the host side, system performance can be degraded in some cases if direct I/Os are requested on an inappropriate data access path. For optimum performance, users have to configure the load balancing policy for different modes depending on the redundancy supported on the disk array side.

Infortrend's Multipathing Solution
Infortrend offers a more flexible multipathing solution that integrates the operation between the host and storage device and better utilizes the special functionality provided by Infortrend's RAID storage subsystems.

For availability, Infortrend supports fail-over redundancy between the host and the EonStor® disk array when any interruption occurs. When configured correctly, disk accessibility will not be interrupted by any hardware failure including HBA or connection link failures.

Regarding performance, Infortrend provides the functionality to optimize bandwidth traffic for both the active/active and active/passive modes in a disk array. The I/O requests will be dispatched to multiple paths according to the load balancing policy. For example, a special driver can check the number of I/O requests on each path and direct I/O requests to the path with lower loading (fewer requests in the queue). This will reduce the possibility that a single I/O path is congested with an overload of I/O requests.

In addition to the traditional native load balancing at the host side, Infortrend also provides an integrated functionality to ensure better performance by cooperating with the EonStor disk arrays. As described above, system performance will most likely degrade if the host distributes I/O requests to an inappropriate path connected to the redundant disk array. Infortrend has an adaptive solution for this occurrence. With the correct configuration, Infortrend's multipathing software on the host side can get the current status of LUNs/disks on the disk array from the RAID controllers and decide the best access path for the specific LUN and direct I/O requests through this path. If there is more than one link in the best path for the specific LUNs/disks, those I/O requests will be balanced to those paths according to the balancing policy.

This special cooperation with Infortrend's storage devices also applies to the Microsoft MPIO architecture. Infortrend provides the specific DSM so that Infortrend disk arrays achieve better availability and improved performance.

Conclusion
By adding a multipathing solution with end-to-end integrated capability to a storage device provider and host platform mechanism, IT managers can ensure continuous, uninterrupted business operations, as well as maximize the efficiency and high availability of computing and storage environments.

Infortrend customers get continuous accessibility to storage with a transparent fail-over mechanism and better performance gains by significantly increasing the available bandwidth with automatic adaptive load balancing to maximize throughput and prevent data path bottlenecks.

This powerful multipathing solution is now available for multiple platforms including Microsoft MPIO and will be extended to other platforms in the near future.