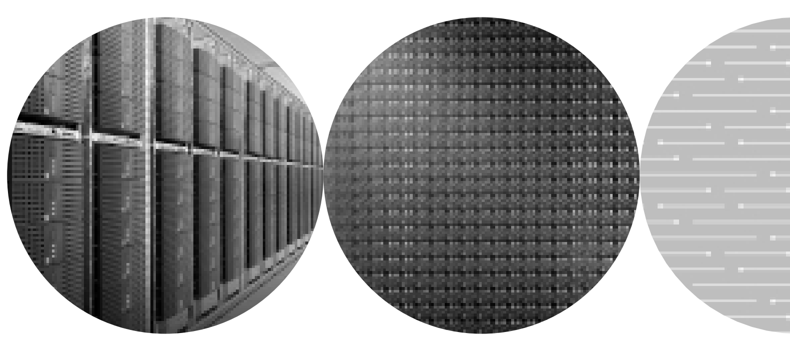
White Paper

Intel in Communications

High Availability Server Clustering Solutions

Extending the benefits of RAID technology into the server arena



Contents

Executive Summary

In today's information technology world, high availability is more and more important in critical systems such as networked storage. Demands on these systems include not only ensuring the availability of important data, but also efficient resource sharing of the relatively expensive components.

By using high availability server clustering in storage environments, organizations can intelligently combine servers and RAID arrays with the purpose of simultaneously increasing the security and the computing power of the whole system. The server cluster appears to the user as one single system with extraordinary characteristics. This concept can also be applied within Storage Area Networks (SANs).

As a strategy to avoid single points of failure in such systems, using host-based RAID controllers for the nodes of the cluster will provide redundancy and security appropriate for mass storage. Intel® Intelligent RAID controllers with high availability clustering support are designed for high-end and enterprise servers, helping to make the benefits of data availability and server reliability accessible to businesses of all sizes.

Extending Protection to Storage Servers

A server cluster, simply put, extends RAID technology into the server arena by creating a system in which one server is redundant to another. This increases the safety and computing power of the entire system.

RAID technology has evolved from the need to simultaneously improve performance, expand capacity, and safeguard information from potential catastrophes such as a power outage or hardware failure. With RAID, several independent hard disks are combined to form one large logical array, or one logical hard disk. Along with data, redundancy information is also stored on the array. This may take the form of mirrored data as in RAID level 1, or parity information as in RAID level 5.

The key is that the operating system no longer deals with separate drives, but with the array as one logical drive. While RAID is not a substitute for backup, since it cannot protect data against theft or fire, it does increase data safety and availability. That protection can be taken a step further. Redundancy can be applied not only to the hard drives, but also to the servers that are attached to the RAID arrays. The two most common means of accomplishing this are mirroring and clustering.

Server Mirroring

Server mirroring simply involves two duplicate servers attached to the same network. One server performs the work while the other receives duplicate (mirrored) data. When one server goes down, the other is ready to take over in a matter of minutes. A major benefit is that the servers can be in separate rooms or buildings for disaster protection. They can even be in different geographic locations if a powerful enough network connection is provided. There are also few hardware restrictions on cabling, controllers or drives.

However, there are significant drawbacks to server mirroring. First, costs can be prohibitive. Organizations purchasing a second server naturally want to put it into production so that it can begin paying for itself, but mirroring requires the second server to be kept in waiting mode. Meanwhile, server maintenance and upgrade costs are effectively doubled.

Performance degradation is another concern. Besides handling all of the file transfer work for network users, the primary server may have to process additional I/Os as it passes information along to the mirror server. This can add substantial processor overhead if system usage is heavy. If there is a partial failure such as the loss of a disk on the primary server and the mirror begins taking up some of the slack, performance will be slower because disk I/O will have to take place over the network. During a full system failure, data may be lost if the primary server cannot complete transfer of in-process data to the mirror before going down.

Server Clustering

Instead of requiring one server to wait on another, server clustering allows each server to act on its own while using one common, redundant mass storage device. In a basic server cluster, two servers would share one RAID array. If one server goes down, the second takes over while still maintaining its own workload. The failover time is greatly reduced because each server is already attached to the information that users are seeking.

Also, there is less chance that data will be lost during a system failure. Data does not have to be sent to a backup system at time of failure, because it is already in both servers. Just as important, the cost factor is reduced since it is not necessary to have a separate RAID array in each server, and the backup server is not merely waiting for something to do. It is functioning as an active production server on the network, sharing the processing load.

Clustering Technologies

Most major operating systems—including those from Microsoft*, Novell*, and Linux providers such as Red Hat* and TurboLinux*—now provide support for server clustering. While details vary, cluster solutions for the most part work in the same basic way.

How Clusters Work

The servers in a cluster communicate with one another through polling, and are configured so that the operating system no longer deals with separate servers, but one logical system (Figure 1). Each server in the cluster continually polls every other server in the cluster to ensure that they are still operational. This polling takes place over a private "heartbeat" network, which requires its own separate cable run.

Each clustered server is connected to:

- The network, via LAN connection
- A common mass storage device, via shared data bus
- The other server(s) in the cluster, through a simple interconnect device such as a secondary network card

When a failover takes place, all three connections become involved. If the failed server is no longer available over the private heartbeat connection, it is polled on the LAN. If this yields no response, then the polling server must take over the failed server's assignment of connecting users to the common data source via the data bus.

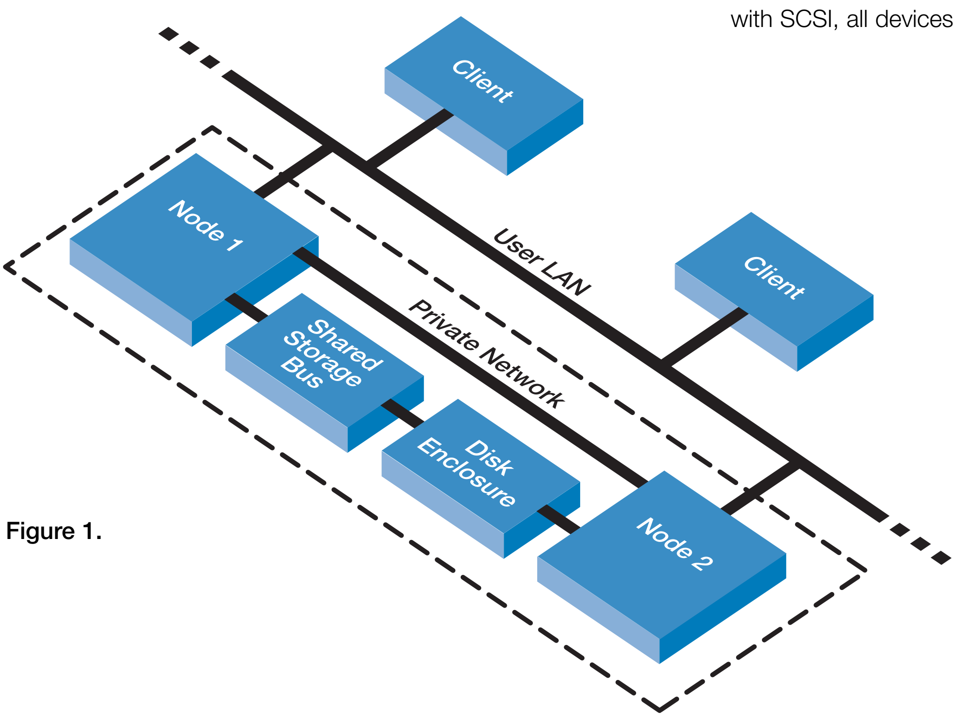


Figure 1.

With interconnected servers operating in a cluster, performance can be improved by adjusting the load on each. With some operating systems, however, dynamic load balancing is not possible at this time. A speed advantage is only possible through static load balancing, where two different applications are running parallel on different servers in the cluster—for example, an SQL Database on one and a Web Server on the other.

Current Connection Types

- SCSI

For servers located on the same rack in the same room, SCSI (Small Computer System Interface) technology is likely sufficient. A parallel data transfer technology, SCSI is economical but has distance and bandwidth limitations.

Even though Ultra 320 has pushed the data transfer rates to 320 MB/s, SCSI still has the limitation of cable length and number of devices. There is also standard differential signaling, which allows up to 25 meters in length, but it is very expensive, not widely used and incompatible with other SCSI formats.

- Fibre Channel

For a higher data transfer rate and longer cable lengths, the choice is Fibre Channel Arbitrated Loop (FC-AL) technology for host-based RAID controllers. The longer cable run allows an organization to increase the number of attached devices per cable and locate them in separate rooms or facilities. As with SCSI, all devices on the cable share the same bandwidth.

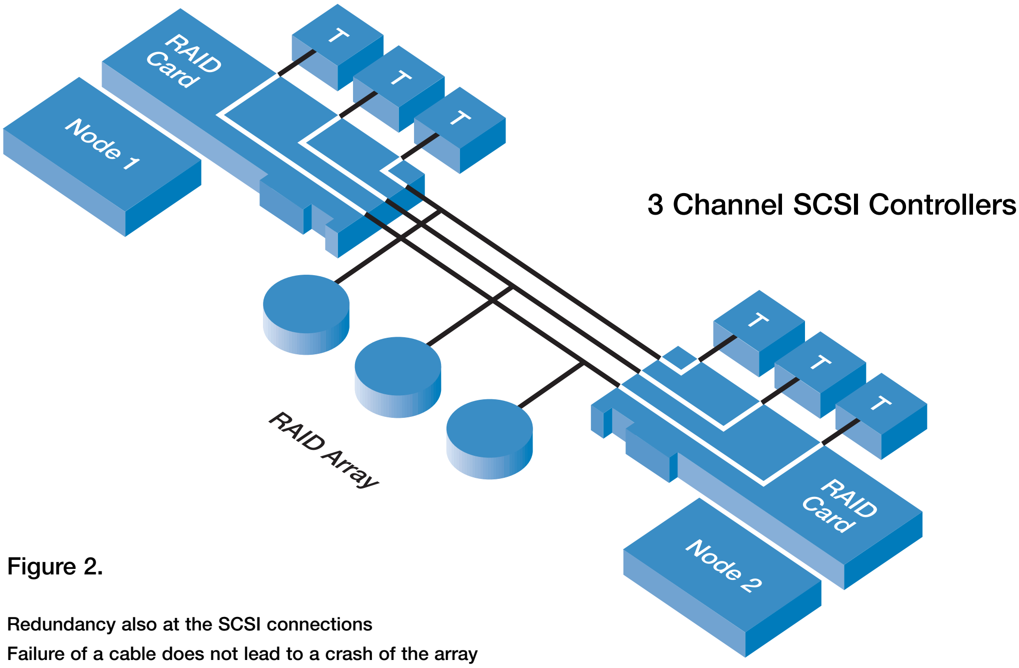


Figure 2.

Redundancy also at the SCSI connections
Failure of a cable does not lead to a crash of the array

Fibre Channel is a serial, high-speed data transfer technology. It is not limited to the transmission of signals through optical fibers, but can also use less expensive copper wiring, either twisted pair or coaxial. FC is an open industry standard and can be viewed as a transport vehicle for the supported command set (usually SCSI commands). Unaware of content, Fibre Channel simply packs data into frames, transports the frames to the appropriate devices, and provides error checking. FC-AL makes it possible to attach up to 126 devices on a single channel and allows distances up to 25 meters using copper cabling. With multi mode optical fiber transmission, a distance of up to 500 meters can be achieved between devices (not just in overall cable length). Using single mode optical fiber, the distance between devices can be up to 10,000 meters or 10km.

Deployment Methods

Although not the only solution, the choice of host-based RAID controllers for the nodes of a cluster helps to eliminate the single point of failure that would be presented by locating control on the disk array. The host-based RAID controller typically resides on an adapter card that offloads RAID-related tasks including host interrupts, freeing the server CPU to handle the application load.

There are several different ways of setting up a cluster. For example, Figure 2 depicts a setup using two cluster servers, each with a three-channel SCSI RAID controller configured in a RAID array. By placing each disk or set of disks on a separate controller, there is also redundancy at the SCSI connections. Therefore, failure of a cable does not lead to a crash of the array.

Using Fibre Channel technology, two dual-channel FC-AL RAID controllers can be used along with redundant Fibre hubs in a dual loop (redundant cable) configuration (Figure 3). In this situation, the Fibre hubs allow the organization to disconnect one server from the cluster and still maintain a certain level of redundancy in the event the enclosure is not configured to automatically close an open loop.

For maximum security, the same configuration could use redundant RAID enclosures (Figure 4).

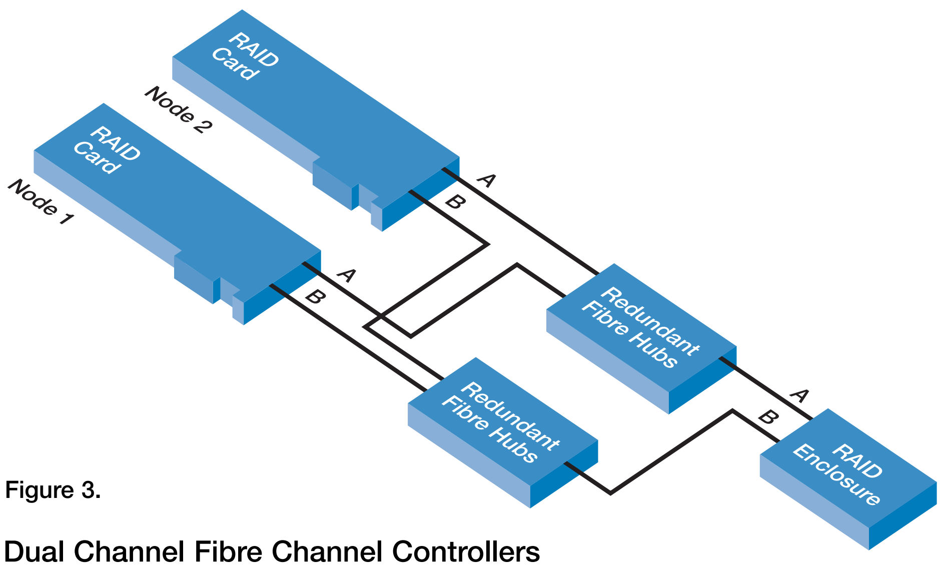


Figure 3.

Dual Channel Fibre Channel Controllers

Fibre Hubs allow to disconnect one node from the cluster Redundant Hubs and "Dual Loop" configuration (redundant cabling)

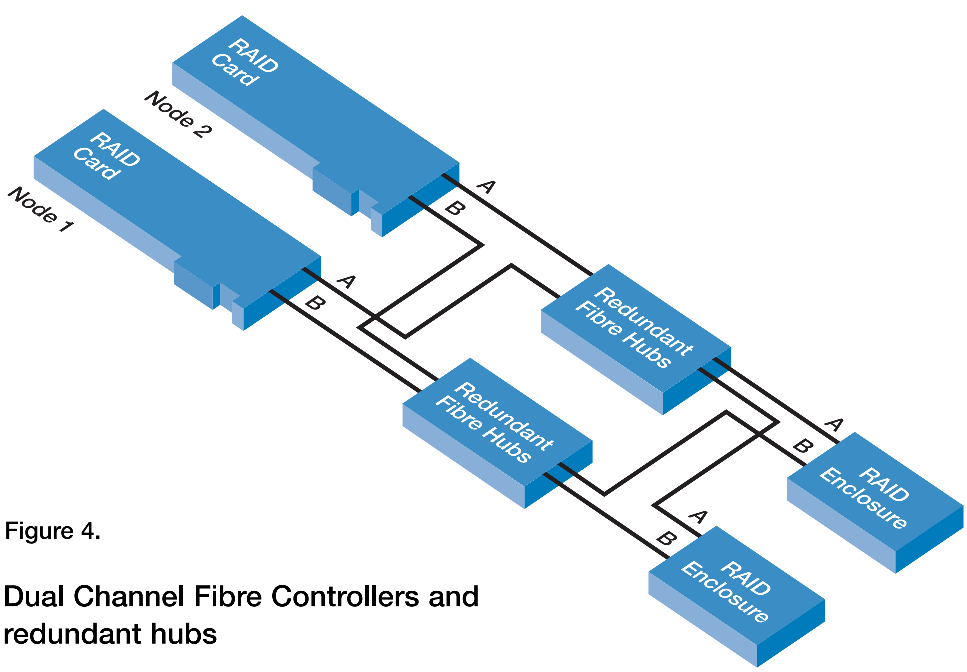


Figure 4.

Dual Channel Fibre Controllers and redundant hubs

Configuration with redundant RAID enclosures for maximum security.

SAN Applications

With increasing requirements for continuous streams of information and zero downtime, many businesses can benefit from a Storage Area Network. However, the cost of implementing a SAN can be prohibitive. The combination of RAID and high availability server clustering can make entry into the SAN storage arena much more feasible. Implementing FC-AL solutions in the SAN environment is relatively easy and can be very powerful.

Fibre Channel is preferable for SANs because it provides outstanding data transfer rates and supports many devices and extremely long cable lengths. Host-based RAID controllers offer an affordable solution into the entry-level SAN field, providing interoperability and scalability.

To set up a very basic SAN "island," an intelligent FC RAID controller is inserted into a server. The server can then be connected to an FC hub that has a link to a disk enclosure. This setup is attached to a LAN via the server. The hub gives this SAN island an easy way to scale into a larger SAN environment when resources permit. For example, adding more storage is as simple as adding another FC disk enclosure to the existing hub.

Additional hubs (or split hub) can carry this one step further by adding cable redundancy and a failover plan. The dual loop configuration of the controller can also add increased bandwidth. Multiple SAN islands can be configured together into a larger SAN environment (Figure 5).

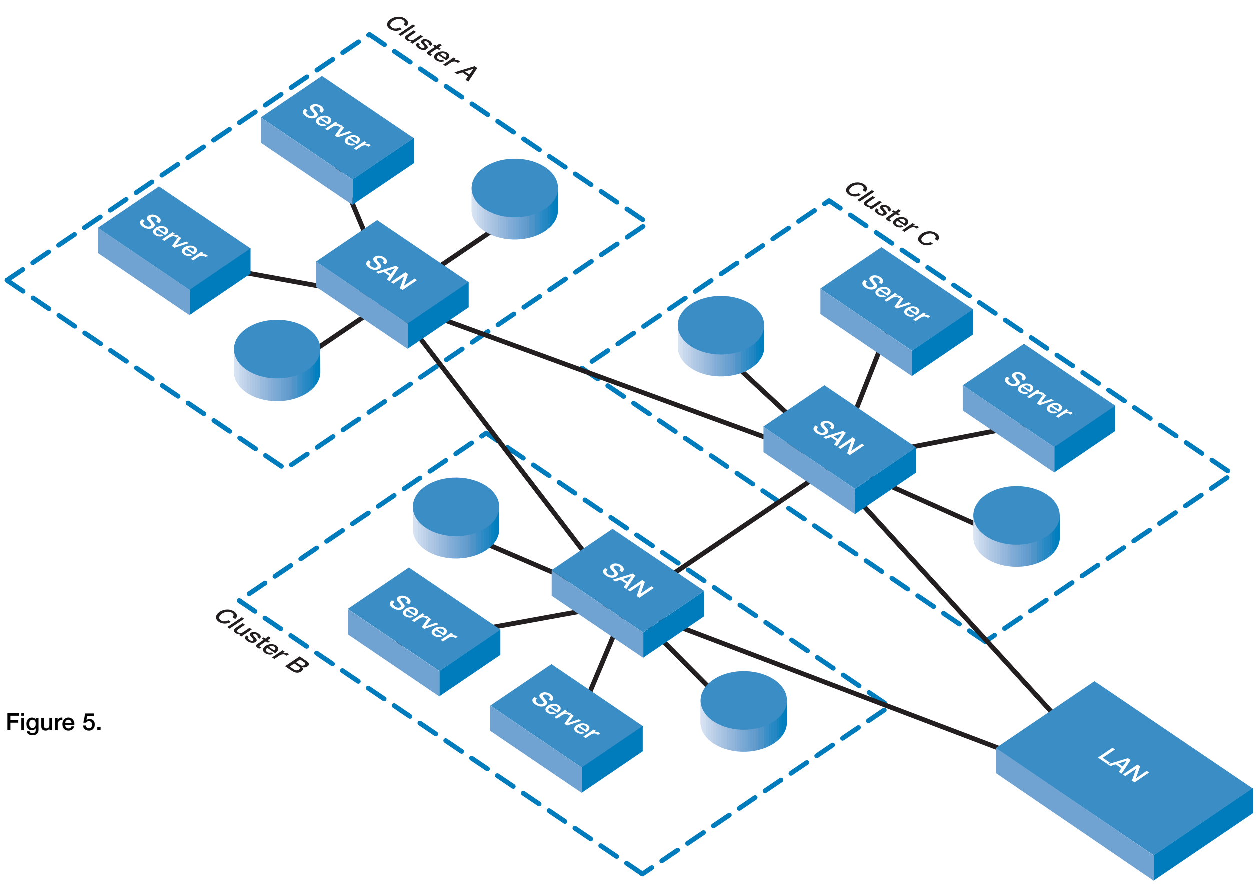


Figure 5.

Conclusion

Increasing network content, database applications, rich email, and e-commerce are driving rapid growth in the volume of data moving across public and enterprise IP networks. In turn, this is driving the demand for affordable, reliable and accessible storage available anywhere on the network.

To meet these needs, any network that runs a mission critical database should consider the combination of RAID and server clustering. This includes online banking sites, auctions, stores or any other type of e-commerce site that depends on uninterrupted access. In fact, the Internet has become a requirement of practically all commerce today. Any downtime can have a severe impact on the entire organization. This is the driving force behind the need for RAID and server clusters as components in more and more enterprise networks.

Intel supplies a wide range of interoperable building blocks for storage at any level of integration, including cluster-capable FC RAID controllers. Intel's open, standards-based approach fosters a community of providers who are accelerating the availability of safe and reliable, high performance storage solutions that result in lower total costs of storage development and ownership.

For more information

For more information about Intel RAID controllers and RAID-on Motherboard, visit:

http://developer.intel.com/design/servers/buildingblocks/