DISK ARRAYS

Table of Contents

1. Definition

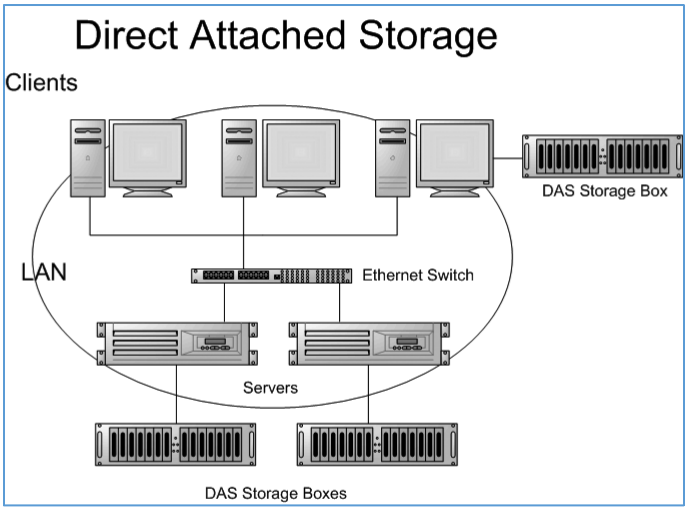
A disk array is a hardware element that contains a large group of hard disk drives (HDDs). It may contain several disk drive trays and has an architecture which improves speed and increases data protection.

The system is run via **a storage controller**, which coordinates activity within the unit. Disk arrays form the backbone of modern storage networking environments. A storage area network (SAN) contains one or more disk arrays that function as the repository for the data which is moved in and out of the SAN.

Typically, disk arrays are divided into categories:

1.1. Direct Attached Storage (DAS)

Direct-attached storage (DAS) refers to a digital storage system directly attached to a server or workstation, without a storage network in between. It is mainly used to differentiate non-networked storage from SAN and NAS.



it is a digital storage system that is directly attached to a server without a storage network in between, for example using an external disk. It is ideal for localized file sharing in small business or departments in work groups. DAS has low scalability and depending on the business long term goals it can increase storage costs in the long run.

1.2. Network attached storage (NAS) arrays

Over 80% of storage capacity sold worldwide will be shipped in support of file-level or ‘unstructured’ data. That is file-based information that doesn’t fit neatly in a column-and-row database. Audio, video, graphics, and all the other content we Facebook and Twitter about are good examples. If you can’t fit all of this neatly in a database, then you need a file system to make order out of the chaos and sit above the low-level disk operations.

NAS is essentially a dedicated appliance with a built-in file system to store these files and then present file shares to servers, applications and users. Most often we think about mapping to these systems using protocols such as SMB/CIFS (for Windows) or NFS (for Linux/Unix). In the past, I’ve heard many people try and plot NAS and SAN on a continuum as if NAS was simply a less sophisticated storage device than a SAN or a stepping stone to SAN, but really, they are completely different animals for completely different purposes. Just as SAN can range from entry-level departmental systems to large tier-1 arrays, the same is true with NAS.

Originally, NAS filers emerged as companies realized that simply putting user shares and file data on random Windows servers or homegrown Linux-based NFS servers was likely not the best way to deal with all this file growth. However, just as people originally were consolidating islands of files into a dedicated NAS filer, they are now looking to consolidate dozens of distributed NAS filers into something easier to manage. This is where scale-out clustered NAS has come to the party. These systems can physically store data on multiple underlying disk technologies but present a large single ‘namespace’ or filesystem out to the network. The concept of a single namespace simplifies management and enables the policy-based movement of files underneath the covers for reasons of performance, capacity, or compliance. IT practitioners love this technology because a system admin can have just-in-time capacity expansion to a large pool of NAS. This provides simplicity in management and access, while avoiding over-provisioning and captive storage.

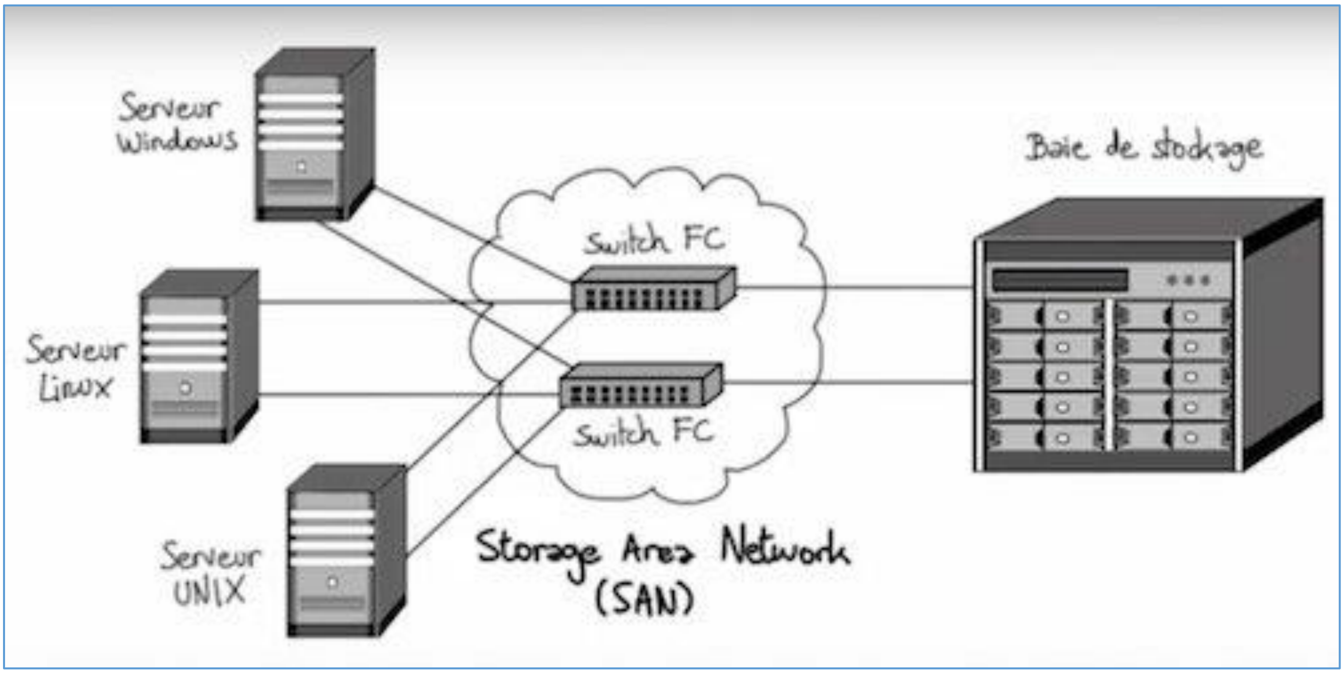
One example of where NAS and SAN come together is the advent of the NAS Gateway. These appliance ‘heads’ are processing nodes with an operating system layer. The NAS head is physically attached to back-end storage capacity that sits in a disk array on the SAN. This gateway manages the filesystem and organization of folders as well as presentation out to the network while the disk array manages the centralized data services we previously discussed. This scenario is a great solution for companies that have invested in a SAN and a disk array but want to increase the utilization of that asset by placing their unstructured files on some of the extra disk. Some might (and do) call this a ‘unified’ storage approach.

1.3. Storage area network (SAN) arrays

- Modular SAN arrays
- Monolithic SAN arrays
- Utility Storage Arrays

The heart and soul of the storage industry continues to be the SAN. (As a note, while in fact the 'SAN' is the actual network to which a disk array is connected, I've found that many people refer to the disk array as 'the SAN'). Despite the meteoric rise in unstructured data growth and file content, SAN arrays still account for most the money spent on external disk storage worldwide. SAN connectivity options traditionally were based around Fibre Channel, but in recent years the broader adoption of 10GbE has firmly planted iSCSI as a mainstream choice. Fibre Channel over Ethernet (FCoE) rounds out the options, but has been a relatively slow starter in the market.

Independent of how you connect, SAN is all about the consolidation of multiple workloads into a shared pool of capacity. The way that capacity is used depends on the host that a given volume/LUN is presented to. The disk array itself simply presents the capacity; any file system or OS requirement is driven by the host on the other side. The disk array itself can provide advanced data services such as multi-site replication and sub-LUN data tiering across different physical disk types. Applying these data services once at the array level can be much more efficient than managing them separately at the host level. This is particularly true in enterprise environments where you may have dozens if not hundreds of physical servers connected to the same disk array.

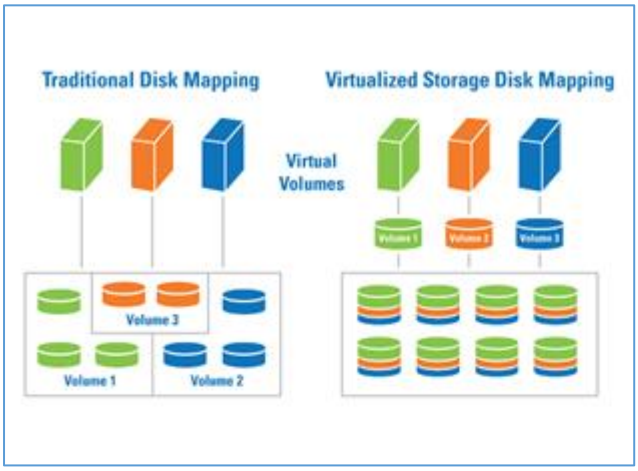


One of the biggest drivers of SAN segment growth in the last 10 years has been virtualization. In addition to utilizing compute resources much more effectively, virtualization has provided the needed abstraction layer between hardware and workloads to provide workload mobility. Moving applications across physical hosts for maintenance, workload balancing, or disaster recovery requires shared storage. It will be interesting to see if hypervisor vendors follow the likes of MS Exchange and build more storage smarts into the hypervisor level. But for now, if you are running virtual servers, chances are that you are connecting to some sort of shared storage device.

One interesting riff on the SAN song and dance is that of the Virtual Storage Appliance–software that can utilize host-based DAS, but pool that capacity together and represent it out to applications as shared capacity. This hybrid doesn't exclusively fit in DAS or SAN, but is really a blending of the two. LeftHand Networks was one of the early pioneers in this space with their VSA solution. Originally positioned as a starter SAN for those not quite ready for dedicated hardware, it has recently found an entirely new market segment as large independent cloud providers look to build out shared infrastructure in a more cost-effective way. VMware has seen the appeal of this approach and recently started including a VSA within its vSphere product. While the VMware VSA is an early-generation product and therefore feature-limited as compared to the LeftHand VSA, it is a good validation of this emerging space.

1.4. Storage virtualisation

Storage Virtualization is the concept of virtualizing enterprise storage at the disk level, creating a dynamic pool of shared storage resources available to all servers, all the time. With read/write operations spread across all drives, multiple requests can be processed in parallel, boosting system performance. This allows users to create hundreds of virtual volumes in seconds to support any virtual server platform and optimize the placement of virtual applications. Storage Virtualization can be structured as Host Based, Storage-Device Based or Network Based. With Storage Virtualization, businesses can create any sized virtual volumes, present network storage to servers simply as disk capacity, automatically restripe data across all drives, dynamically scale the resource pool and use virtual ports to increase port capacity, disk bandwidth, I/O connectivity and port failover.



2. Disk Array Vendors

Primary vendors of storage systems include EMC Corporation, Hitachi Data Systems, NetApp, IBM, Hewlett-Packard, Oracle Corporation, Dell, Infortrend, Coraid, Inc., Panasas and other companies which often act as OEMs for the above vendors and do not themselves market the storage components they manufacture.

EMC established itself in the 1990s as a disk array specialist. However, there are many other vendors offering disk arrays, including NetApp, IBM, Dell, Oracle, HP, Hitachi Data Systems, Compellent, Pillar Data Systems, Xiotech, Fujitsu, NEC, Data Direct Networks, Huawei Symantec, Nexsan, SGI, Overland and Infotrend, among others.

At the lower end of the market, there are a wide range of vendors offering relatively cheap disk arrays, including Buffalo, LaCie, Dulce Systems, G-Technology, Newer Technology and EMC's Iomega division. Vendors such as Dell, HP and Overland also offer inexpensive lower-end disk arrays.

3. LUNs and RAID

Logical Units: The data within a disk array is organized into Logical Units (LUs). Small Computer Systems Interface (SCSI) I/O commands are sent to a storage target and executed by an LU within that target.

LUN: A Logical Unit Number (LUN) is a unique identifier that is used to distinguish between devices that share the same SCSI bus. Commands that are sent to the SCSI controller identify devices based on their LUNs.

Redundant Array of Inexpensive (or Independent) Disks (RAID): RAID employs two or more drives to improve performance and fault tolerance. RAID enables the storage of data in multiple paces to protect the data against corruption and to serve it to users faster. There are various levels of RAID. RAID 0, for example, provides data striping, whereby data is spread out across many disk drives without having any redundancy. This is a good way of improving performance. RAID 1, on the other hand, has disk mirroring so that all data is written to two physical disks and the complete data exists in two places; if one disk fails, the other one is available. This is a good way to speed reads, but the downside is that it slows the write process. RAID 10 is a combination of RAID 0 and 1. There are also further types of RAID, like RAID 5 and RAID 6.

4. Disk Array Technology

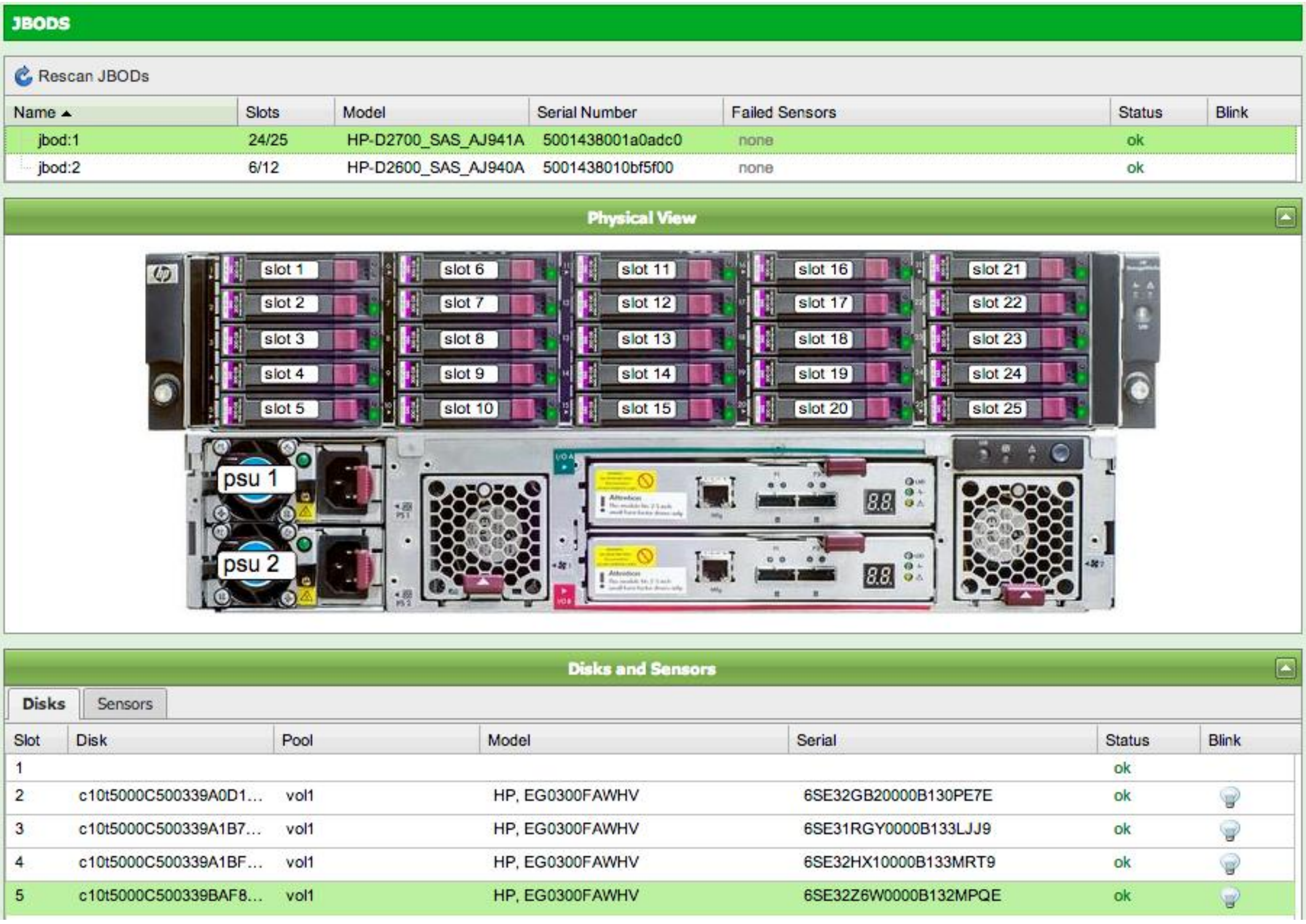
Disk arrays contain the many spinning disks where data is stored. In other words, a whole series of disks are arrayed inside. In such a grouping, many physical disks can be combined to form much larger **logical disks**. The storage controller can dictate that three physical disks of 100 GB be combined into a logical volume of 300 GB, for example. When you have hundreds of spinning disks in the unit, you can then create a few huge logical volumes if desired.

The term disk array also applies to a wide range of different devices. At the high end, monolithic storage arrays are very large and expensive and come complete with enterprise-class features such as full redundancy. If one disk or other component fails, the system continues without pause. Thus enterprise-class disk arrays are designed to be highly available. They hardly ever go down. Further, they come with a large quantity of memory. Frequently accessed data can be stored in memory and served rapidly.

Enterprise disk arrays also make heavy usage of RAID. These arrays often contain two high-performance redundant RAID controllers to serve their large quantity of disks. Some RAID configurations mean that the system remains running if one disk fails. However, RAID 6 is an example of a type of RAID which can cope with two simultaneous disk failures without downtime.

In addition, there are a wide range of disk arrays to serve the midrange, small business and even the consumer markets. These days it is even possible to purchase multi-TB disk arrays for a few hundred dollars. However, these products lack sophistication, don't come with the redundancy and other features of enterprise arrays, and are performance constrained. Whereas an enterprise system might be able to serve data efficiently to mainframe computers processing a dizzying amount of transactions per second, low-end disk arrays are mainly aimed at environments where only a few PCs or servers are accessing the data.

Over time, even the low-end boxes are incorporating RAID and basic data protection technologies into their arsenal. And as processor and connection speeds increase, they become more and more capable of achieving a reasonable level of performance.

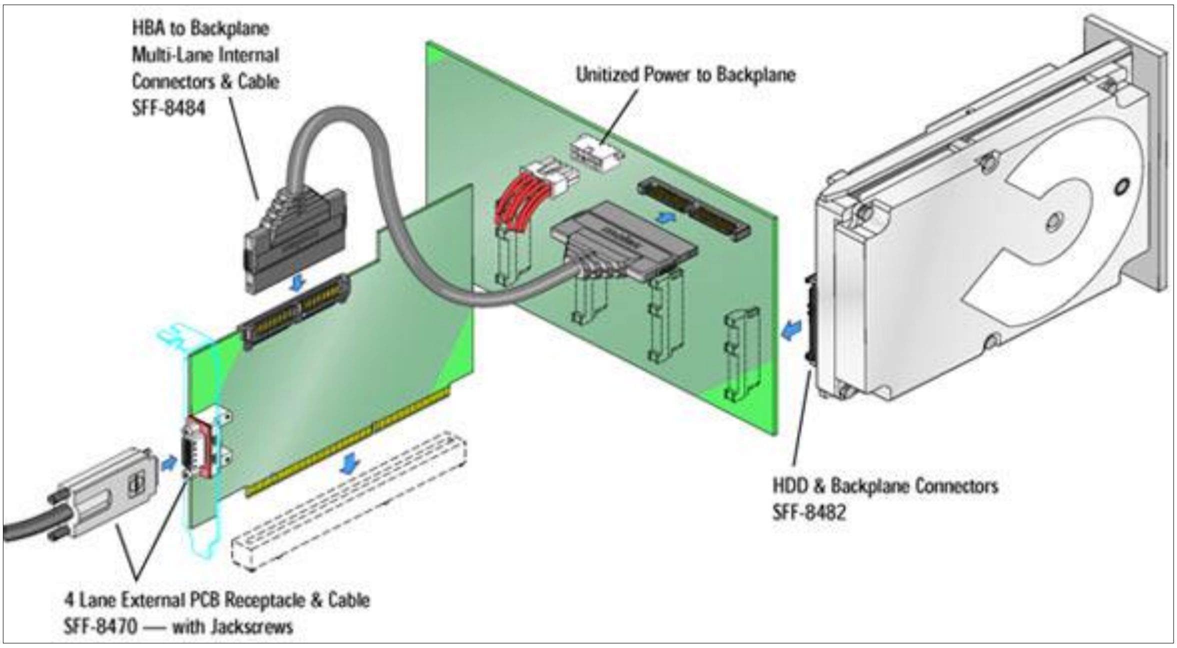


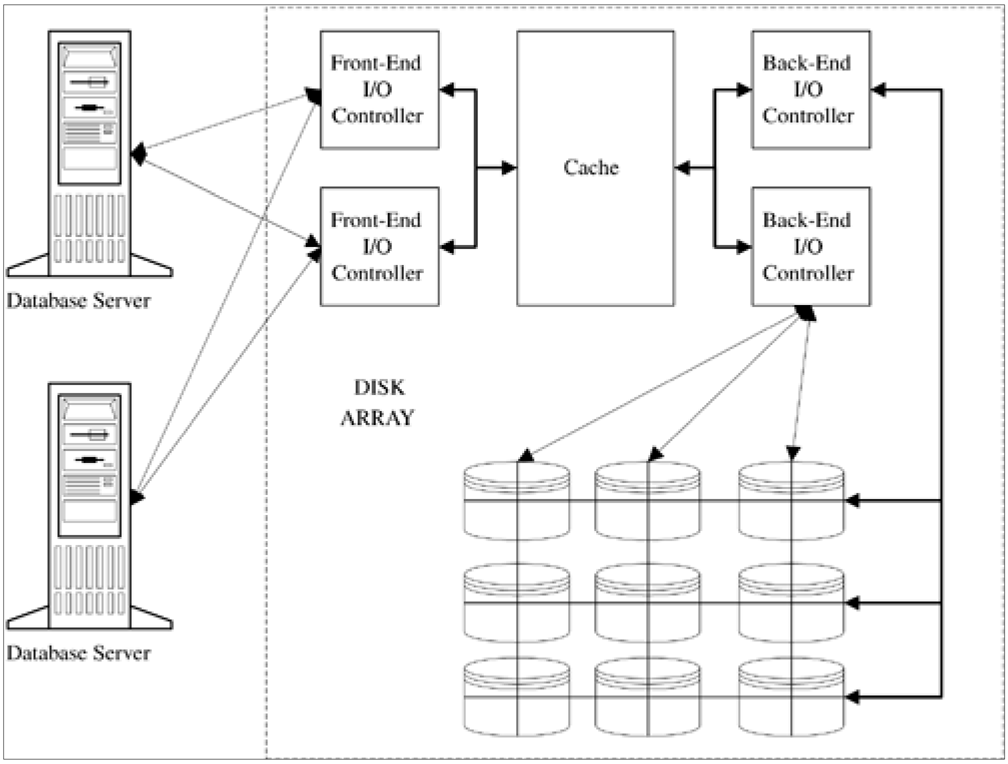
4.1. Inside a Disk Array

Backplanes in storage

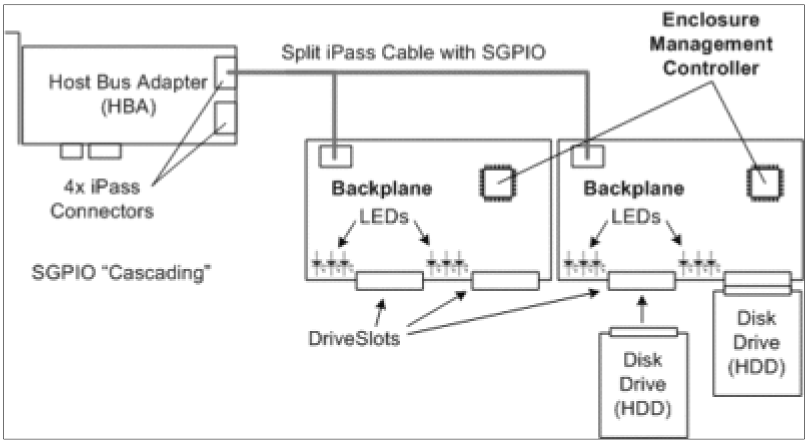
Servers commonly have a backplane to attach hot swappable hard drives; backplane pins pass directly into hard drive sockets without cables. They may have single connector to connect one disk array controller or multiple connectors that can be connected to one or more controllers in arbitrary way. Backplanes are commonly found in disk enclosures, disk arrays, and servers.

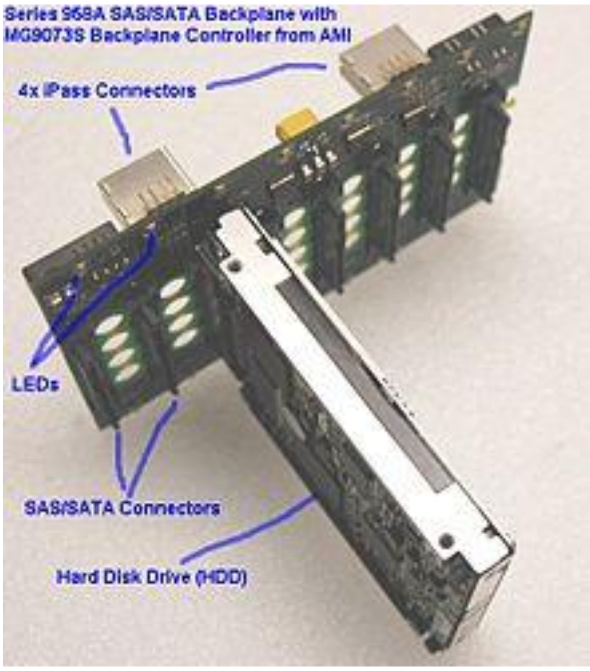
Backplanes for SAS and SATA HDDs most commonly use the SGPIO protocol as means of communication between the host adapter and the backplane. Alternatively, SCSI Enclosure Services can be used. With Parallel SCSI subsystems, SAF-TE is used.





A **backplane** is a circuit board with connectors and power circuitry into which hard drives are attached; they can have multiple slots, each of which can be populated with a hard drive. Typically, the backplane is equipped with LEDs which by their colour and activity, indicate the slot's status; typically, a slot's LED will emit a particular colour or blink pattern to indicate its current status.


Enclosure
Management
Controller
Split iPass Cable with SGPIO
Host Bus Adapter
(HBA)
4x iPass
Connectors
SGPIO "Cascading"
Backplane
LEDs
Backplane
LEDs
DriveSlots
Disk
Drive
(HDD)
Disk
Drive
(HDD)


Series S68A SAS/SATA Backplane with
MG9073S Backplane Controller from AMI
4x iPass Connectors
LEDs
SAS/SATA Connectors
Hard Disk Drive (HDD)