

.....**SHARED DISK FILE SYSTEMS**

The greatest performance limitation of NAS servers and self-configured file servers is that each file must pass through the internal buses of the file servers twice before the files arrive at the computer where they are required (Figure 4.7). Even DAFS and its alternatives like NFS over RDMA cannot get around this ‘eye of the needle’. With storage networks it is possible for several computers to access a storage device simultaneously. The I/O bottleneck in the file server can be circumvented if all clients fetch the files from the disk directly via the storage network (Figure 4.11).

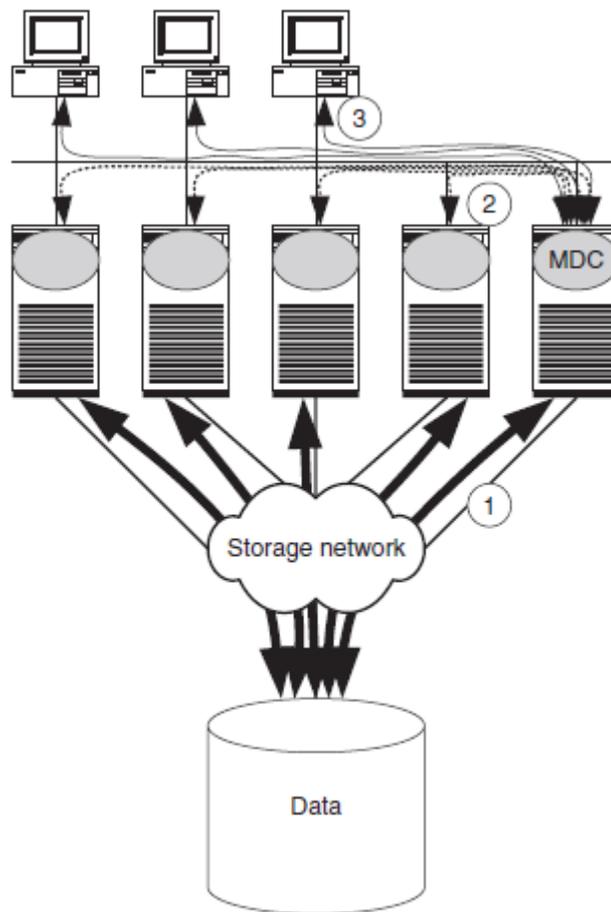


Figure 4.11 In a shared disk file system all clients can access the disks directly via the storage network (1). LAN data traffic is now only necessary for the synchronisation of the write accesses (2). The data of a shared disk file system can additionally be exported over the LAN in the form of a network file system with NFS or CIFS (3).

The difficulty here: today's file systems consider their storage devices as local. They concentrate upon the caching and the aggregation of I/O operations; they increase performance by reducing the number of disk accesses needed. So-called shared disk file systems can deal with this problem. Integrated into them are special algorithms that synchronise the simultaneous accesses of several computers to common disks. As a result, shared disk file systems make it possible for several computers to access files simultaneously without causing version conflict.

To achieve this, shared disk file systems must synchronise write accesses in addition to the functions of local file systems. It should be ensured locally that new files are written to different areas of the hard disk. It must also be ensured that cache entries are marked as invalid. Let us assume that two computers each have a file in their local cache and one of the computers changes the file. If the second computer subsequently reads the file again it may not take the now invalid copy from the cache.

The great advantage of shared disk file systems is that the computers accessing files and the storage devices in question now communicate with each other directly. The diversion via a central file server, which represents the bottleneck in conventional network file systems and also in DAFS and RDMA-enabled NFS, is no longer necessary.

In addition, the load on the CPU in the accessing machine is reduced because communication via Fibre Channel places less of a load on the processor than communication via IP and Ethernet. The sequential access to large files can thus more than make up for the extra cost for access synchronisation. On the other hand, in applications with many small files or in the case of many random accesses within the same file, we should check whether the use of a shared disk file system is really worthwhile.

One side-effect of file sharing over the storage network is that the availability of the shared disk file system can be better than that of conventional network file systems. This is because a central file server is no longer needed. If a machine in the shared disk file system cluster fails, then the other machines can carry on working. This means that the availability of the underlying storage devices largely determines the availability of shared disk file systems.