



Understanding Media Pools

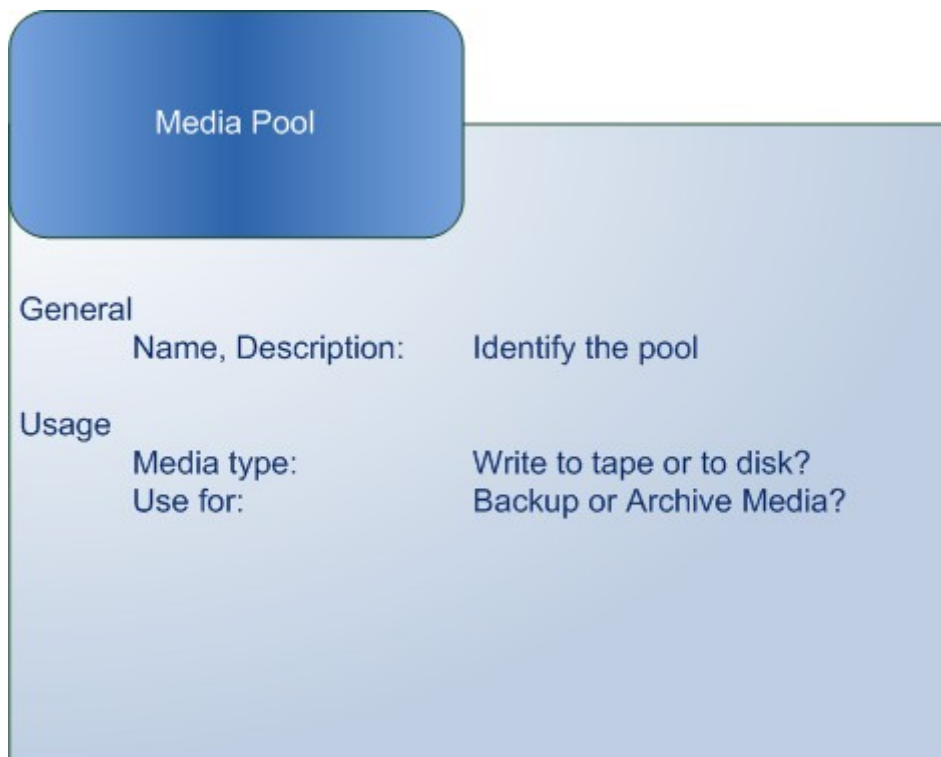
This article targets the PresSTORE backup power users. It offers a closer look to the organization and mechanisms given by the Media Pools.

Why does PresSTORE use Media Pools?

When having a single tape device and a bookshelf with tapes, Media Pools actually would not be required. The “bookshelf organization” does not need a computer. But already when using a small tape library with two tape magazines, a little bit of organization is required: Daily backup on one side, weekly on the other. Running with multiple storage devices, e.g. a single tape drive for archiving, a juke box for backup and a hard disk for daily backups requires proper administration .

The Basics

A Media Pool acts as a container for media. When tapes (media) are labeled, i.e. when they are prepared for the use with PresSTORE, they are assigned to a Media Pool. Thus each pool makes up a set of tapes. It has a name and a purpose that it is used for:



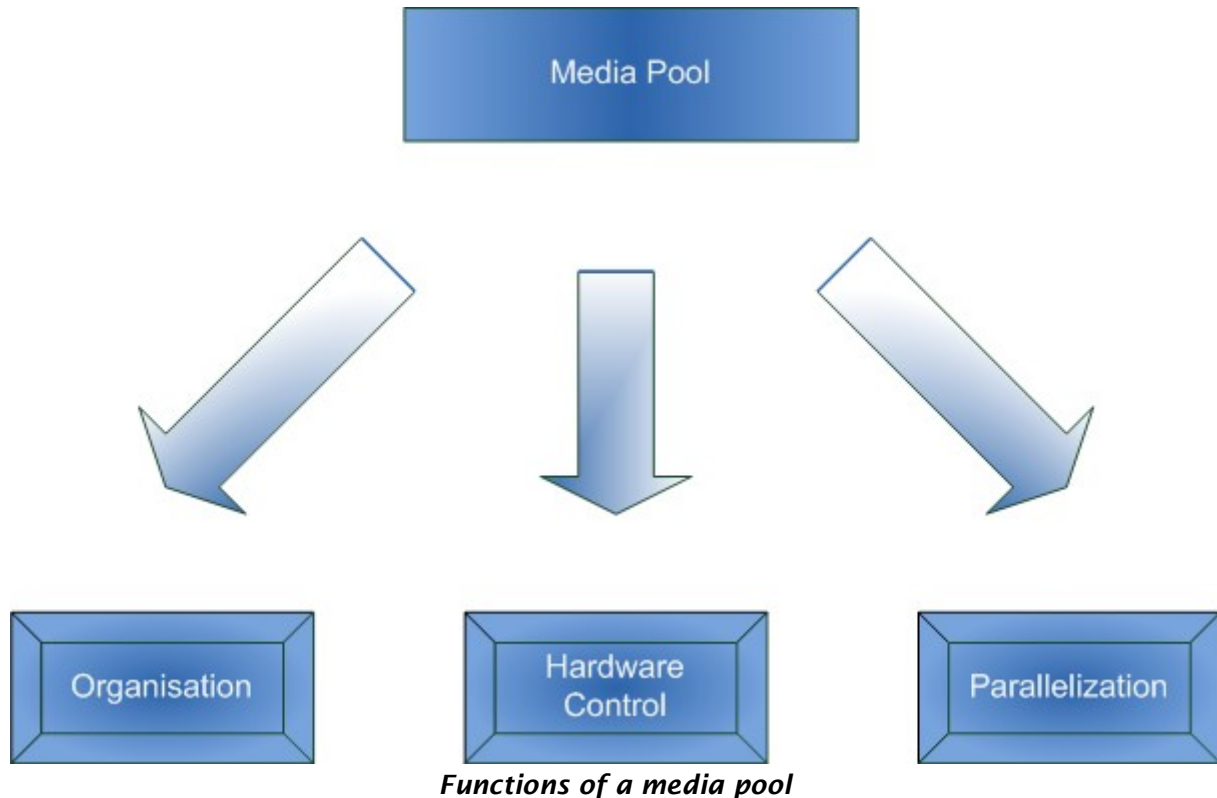
Basic Media Pool properties

As a result, each Media Pool stands for a set of tapes, foreseen for a specific purpose. So the pool is a potential storage, limited by the number of tapes it contains, but expandable by labeling further tapes.



The Functions

Media Pools serve, besides organization, to control hardware resources (namely drives) and control the parallelization of read and write operations.



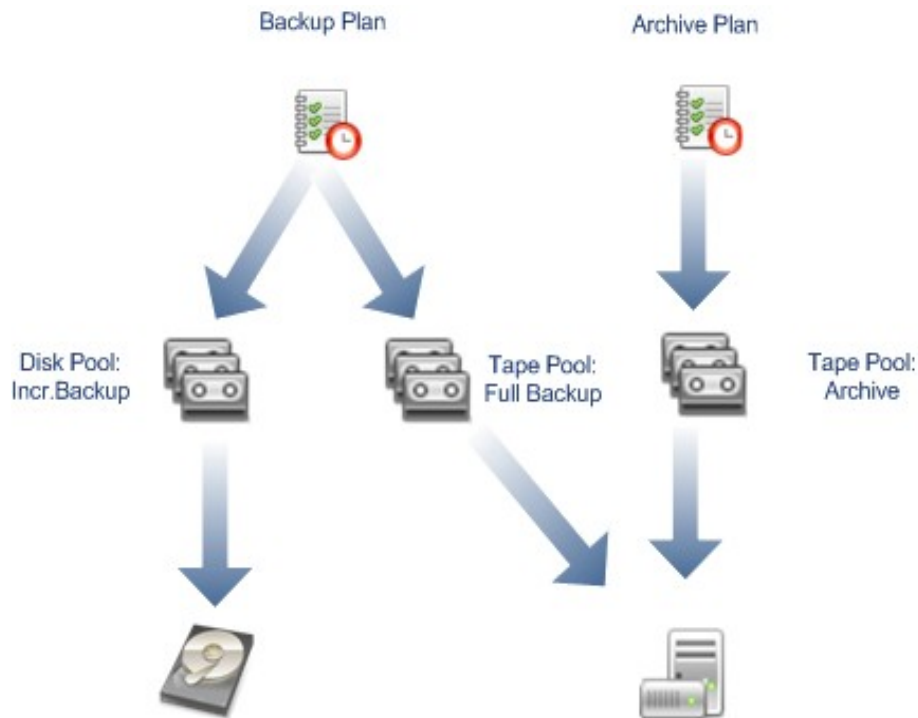
Media Organization

When PresSTORE Backup or Archive plans are defined, they do not refer to tapes but to a Media Pool. Where the plan organizes content, time schedule and repetitions, the Media Pool organizes media and hardware access.

It also helps separating different sets of tapes and simplifies the handling, as different Pools may be for example located in different magazines or trays (or slot areas, if trays are not contained in the juke box). It is not required that all tapes of a pool reside in one magazine, for example some tapes of the pool may be offline; but it is helpful to know the tape's position in the jukebox.

Tape Handling

Media Pools are extremely useful when organizing tape handling. By example, let us define a full backup to tape with incremental backups to disk and from time to time some files are archived:



Jobs and pools, a configuration example

We will need a disk pool for incremental backups. This is a separate organization unit using a virtual jukebox on disk with its own virtual tapes and recycling strategy.

In the real jukebox, two kinds of tapes are required: those for archive, never to be overwritten, and those for backup, with automated recycling. The backup tapes are furthermore exchanged in regular intervals, so an easy way is needed to separate the archive tapes from the backup tapes. To cover this, just create two tape pools, one for archive, one for backup. The archive tapes may reside in one magazine, the backup tapes in another.

There is a convention always adhered to in PresSTORE: each tape is always kept in the same slot where it was first seen. Because of this, the archive tapes and the backup tapes are not mixed up. When exchanging backup tapes, one can simply exchange the cartridge with the backup tapes without thinking of the archive tapes.

Recycling of Backup Tapes

While archive tapes are once written and then kept as long as they live, backup tapes are reused after a while when the backup is renewed. There are two strategies that can be used:

- Recycling as soon as possible, that is when the recycling date is reached. This means that recyclable tapes are preferred to empty tapes. So when the backup requires less tapes than available, some tapes remain empty.



- Recycling when tapes are needed is the opposite. This way, tapes are recycled only if there is no other way. When the backup requires fewer tapes than available, this strategy keeps additional old backup tapes as long as possible.

Hardware Control

Hardware control is possible in multiple ways. All together the Media Pool settings built the frame in which the jobs will later be executed.

Drive Assignment

Each media pool can be restricted to specific drives, that means when writing to the pool, only these drives are used. This restriction is mostly not required, but there are some situations where limiting the drives can be helpful:

- Your jukebox has an LTO2 and an LTO3 tape drive, and you want to restrict the drive usage to the appropriate tapes. This can be achieved by creating two media pools, each one assigned to one of the drives. Just label your tapes for the according Media Pool.
- Your jukebox has two drives, but you want the daily backup to keep one drive free for possible restore jobs: Just assign one drive to your Media Pool. (Something similar can be achieved by restricting the parallelism, see below).
- You have two disks with virtual jukeboxes and want to use them in parallel. To do that, enhance the parallelism to allow two drives (this is explained later), and in addition to that restrict the pool to use one drive of each jukebox. Otherwise, the two drives of one jukeboxes are selected.

Cloning

The Media Pool also controls the duplication of tapes, which is also called cloning. For each tape written in drive 1, a second identical tape is written in another drive. While doing this, PresSTORE ensures that both tapes contain identical data, so these tapes can be reconstructed from the clone in case one of the two tapes is damaged.

Cloning is recommended for archives. In case only a single tape contains the archived data, a damage of that tapes will destroy all data. With an existing clone, the data can be reconstructed.

As cloning requires identical data, the clone must be written at the same time as the original tape. Therefore, a second tape drive is required to enable cloning.



Jukebox and Drive Control

Depending on the Media Pool, it can be configured whether to keep the tape in a drive or whether to move it back to the slot in the jukebox where it came from. The latter is the default.

The first option can be helpful to speed up smaller jobs, e.g. for archiving. The archive tape(s) may be left in the drive(s), only during a backup they are moved back to their slots, same as the backup tapes.

For single tape drives, it can be defined whether the mount procedure shall be done manually or automatically. When turning off the automount option, PresSTORE will not become aware of changed tapes (tapes then must be mounted manually). Turning off the auto mount option can be helpful when a backup tape becomes full and it is unwanted that PresSTORE checks every few minutes whether there is a new tape and reads the full tape for that purpose, for example in full backups over the weekend.

Parallelizing

The usual goal when parallelizing jobs is to enhance the degree of efficiency and get things done faster. As this theme is too complex to fit into a single paragraph, a separate article illuminates the speed optimization with parallelisms:

[Parallelizing for optimization, Archiware, 03/2008]

By default, the parallelism is set to one data stream with an unlimited number of clients and one drive, this means that by default all Media Pools make use of only one drive at a time (except where cloning is involved).

Two items often used are worth to give them a closer look:

- In order to allow the parallel use of two drives, navigate to the Parallelizing Options of the Media Pool and increase the drives per stream to 2. This way, all jobs writing to this pool will use two drives. In conjunction with the above mentioned drive assignment, this gives a possibility to canalize the data flow.
- When multiple clients are to be backed up, the network capacity may become a bottleneck. E.g. when using a 1 GB Network and saving data to one LTO3 drive, a number of four clients is a reasonable number of clients to save in parallel. More clients will not offer a speed advantage, less clients will make the tape operate in a stop and go mode. Reducing the number of clients per stream e.g. to four and limiting the number of streams to one means that the first four clients are backed up simultaneously, the next four when the first four are done, and so on.

Media Pools and Virtual Jukeboxes

As disk storage is getting cheaper every day, and a 1 TB hard disk costs less than a tape



drive, many backup systems save to disk instead of to tape. PresSTORE support this by creating so called virtual jukeboxes on hard disks. These are fully functional software emulations of real media changers.

All mechanisms described above for tapes can be applied for disk backups, too. The only exception is that the virtual tapes, represented by disk files, are usually never removed from the virtual jukebox.