

Multiple Mount Protection

Multiple mount protection (MMP) is a feature that protects the filesystem against multiple hosts trying to use the filesystem simultaneously. When a filesystem is opened (for mounting, or fsck, etc.), the MMP code running on the node (call it node A) checks a sequence number. If the sequence number is EXT4_MMP_SEQ_CLEAN, the open continues. If the sequence number is EXT4_MMP_SEQ_FSCK, then fsck is (hopefully) running, and open fails immediately. Otherwise, the open code will wait for twice the specified MMP check interval and check the sequence number again. If the sequence number has changed, then the filesystem is active on another machine and the open fails. If the MMP code passes all of those checks, a new MMP sequence number is generated and written to the MMP block, and the mount proceeds.

While the filesystem is live, the kernel sets up a timer to re-check the MMP block at the specified MMP check interval. To perform the re-check, the MMP sequence number is re-read; if it does not match the in-memory MMP sequence number, then another node (node B) has mounted the filesystem, and node A remounts the filesystem read-only. If the sequence numbers match, the sequence number is incremented both in memory and on disk, and the re-check is complete.

The hostname and device filename are written into the MMP block whenever an open operation succeeds. The MMP code does not use these values; they are provided purely for informational purposes.

The checksum is calculated against the FS UUID and the MMP structure. The MMP structure (`struct mmp_struct`) is as follows:

| Offset | Type | Name | Description |
|--------|-------------|--------------------|--|
| 0x0 | __le32 | mmp_magic | Magic number for MMP, 0x004D4D50 ("MMP"). |
| 0x4 | __le32 | mmp_seq | Sequence number, updated periodically. |
| 0x8 | __le64 | mmp_time | Time that the MMP block was last updated. |
| 0x10 | char[64] | mmp_nodename | Hostname of the node that opened the filesystem. |
| 0x50 | char[32] | mmp_bdevname | Block device name of the filesystem. |
| 0x70 | __le16 | mmp_check_interval | The MMP re-check interval, in seconds. |
| 0x72 | __le16 | mmp_pad1 | Zero. |
| 0x74 | __le32[226] | mmp_pad2 | Zero. |
| 0x3FC | __le32 | mmp_checksum | Checksum of the MMP block. |