

# Linear Tape File System

**David Pease**  
**IBM Almaden Research Center**

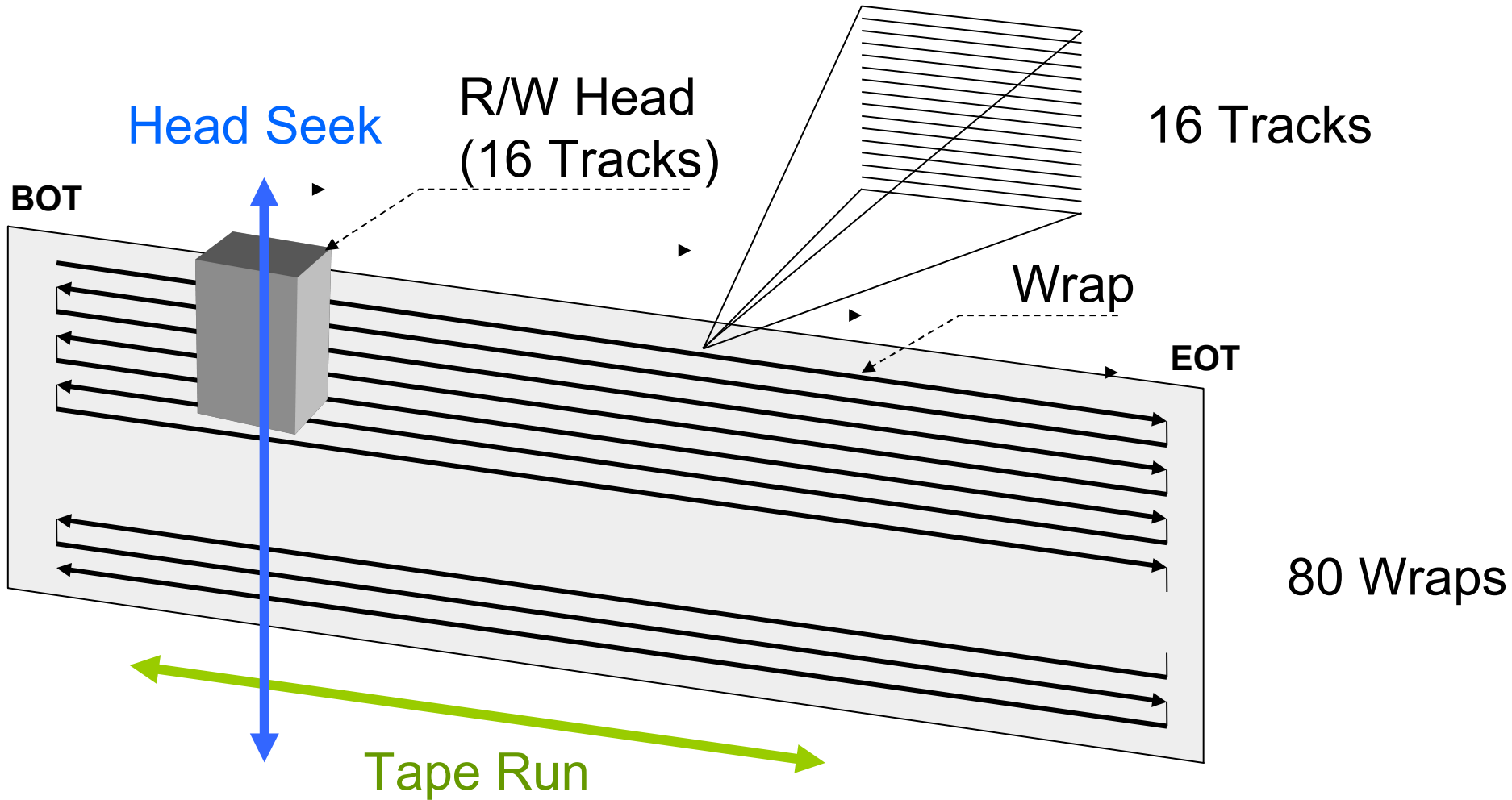
# Tape Storage Today

- Tape is very much alive
  - \_ 2008 study estimated 51% of all archive data is on tape
    - 5 exabytes, expected to grow to 24 exabytes by 2012
  - \_ Tape can have many advantages over disk
    - one study estimates cost ratio for terabyte stored long-term on SATA disk versus LTO tape is about 23:1
    - energy cost ratio is as high as 290:1
    - bit error rate of a SATA hard drive is at least an order of magnitude higher than of LTO tape
    - longevity is typically rated at 30 years
    - high capacity and data streaming rate
- However, tape is often not convenient to use
  - \_ No real standard format for data on tape
    - tar is closest, has many drawbacks
  - \_ Tapes often managed by external manager (e.g. TSM)
    - Relies on external database, tapes not self-contained
  - \_ Not considered a good medium for data interchange

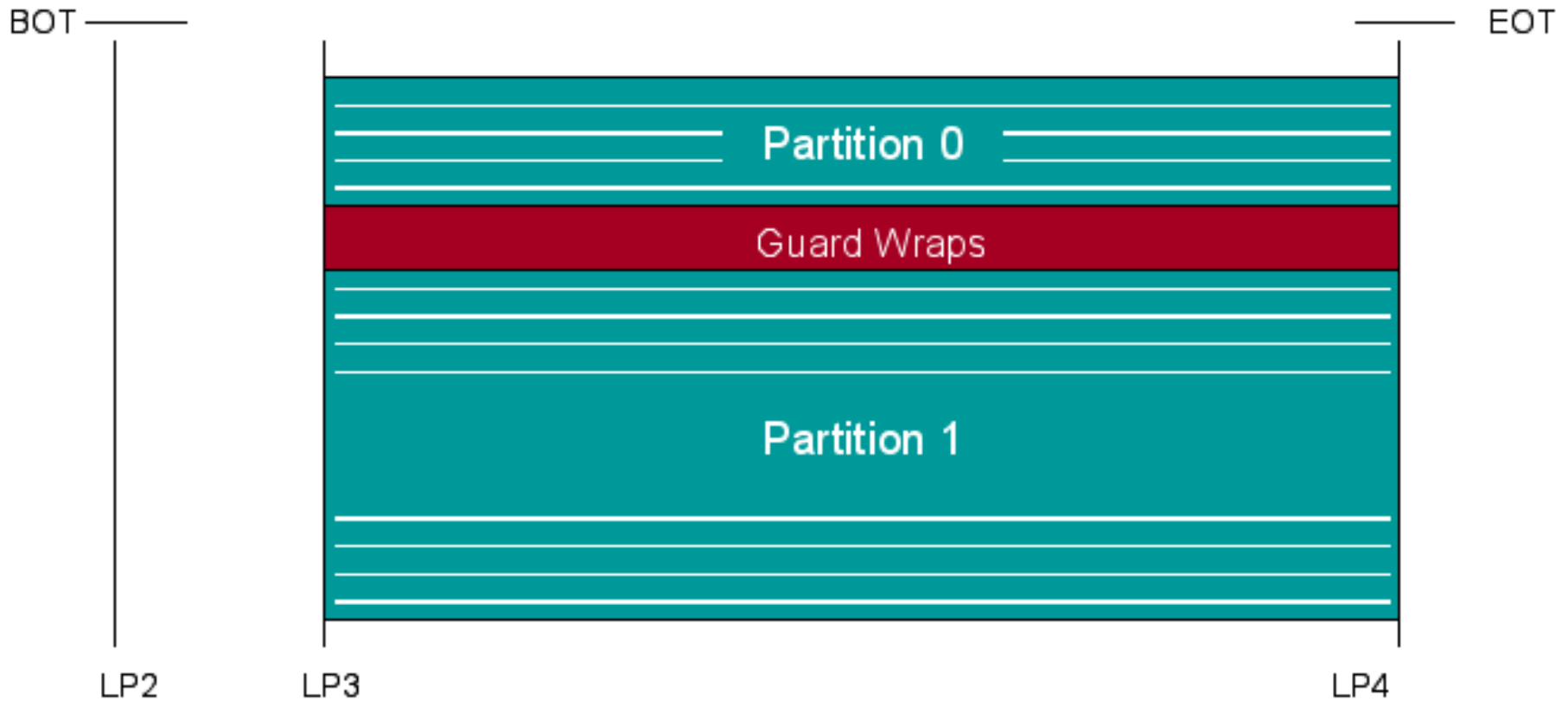
# Linear Tape Open (LTO) Tape

- LTO Consortium
  - Defines an industry standard for tape drives and media
    - IBM, HP, Quantum, many media manufacturers
- LTO Tape
  - Serpentine recording, shingled writing
  - Block-addressable
  - Essentially an append-only media
- LTO Generation 5 (LTO-5)
  - Released April 2010
  - 1.5 terabytes per cartridge (uncompressed)
  - 140 MB/sec streaming data rate
  - Dual-partition capability

# LTO Track Recording



# Dual-Partition Tape – logical view



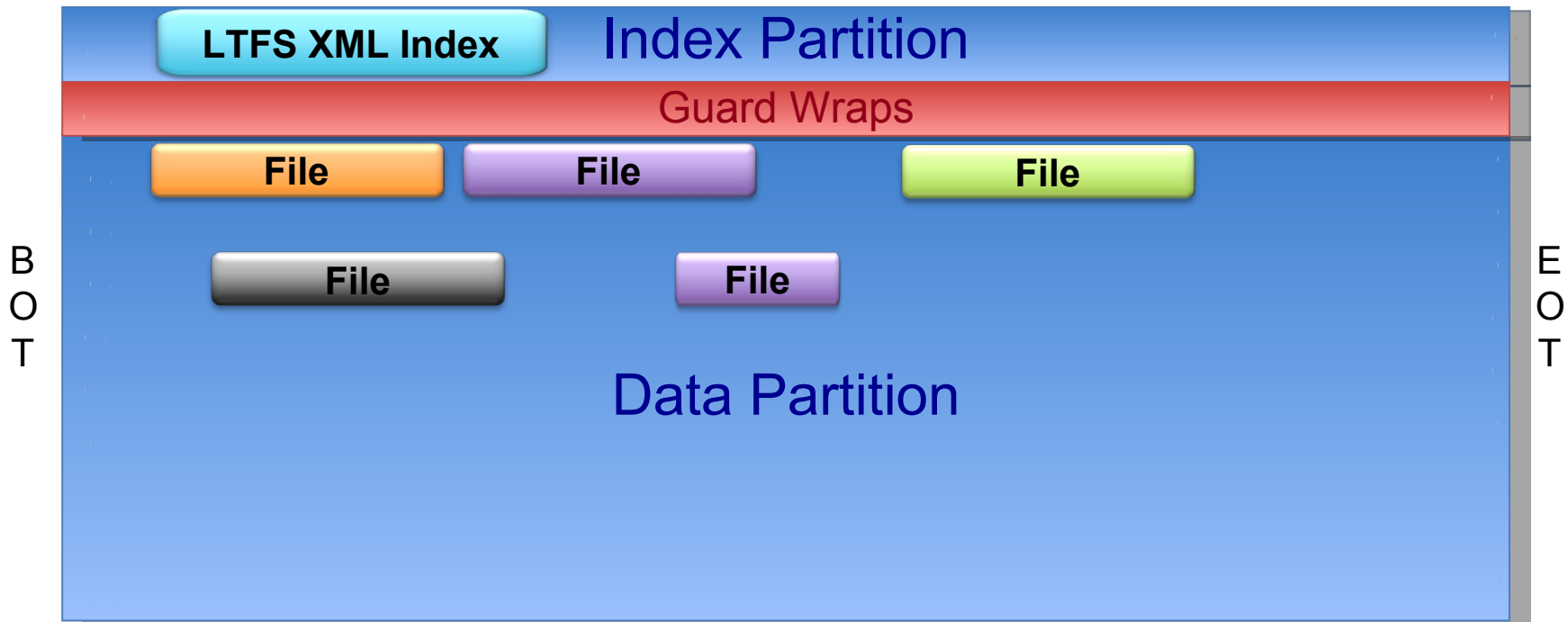
# What is LTFS?

- A file system implemented on dual-partition linear tape
  - Makes tape look and work like any removable media (e.g., USB drive, removable disk)
    - Files and directories show up on desktop, directory listings, etc.
    - Drag-and-drop files to/from tape, double-click to open
    - Run any application written to use disk files
  - Supports automated libraries as well as stand-alone drives
    - In library mode, allows listing contents and searching of all volumes in library without mounting tapes
  - IBM single-drive implementations released as open source
    - Linux and Mac OS X versions released as Open Source
    - Windows version freely downloadable
    - [www-03.ibm.com/systems/storage/tape/ltfs/index.html](http://www-03.ibm.com/systems/storage/tape/ltfs/index.html)

# What is LTFS?

- A file system implemented on dual-partition linear tape:
  - Index Partition and Data Partition
    - Index Partition is “small” (2 wraps, 37.5 GB)
    - Data Partition is remainder of the tape
  - File System module that implements a set of standard file system interfaces
    - Implemented using FUSE
      - On Linux and Mac OS X
    - Windows implementation uses FUSE-like framework
  - Includes an on-tape structure used to track tape contents
    - XML Index Schema

# Logical View of LTFS Volume





# XML Index Schema

- Similar to information in disk-based file system
  - Files
    - Name, dates, extent pointers, extended attributes, etc.
  - Directories
- Designed to be simple, cross-platform
  - Tags and values easy to read, “human” format
  - No platform-specific data
    - Supports Unix/Linux, MacOS, and Windows
- We expect this format to become standard for linear tape
- Format specification on LTO Consortium site:
  - [www.trustlto.com/LTFS\\_Format\\_To\\_Print.pdf](http://www.trustlto.com/LTFS_Format_To_Print.pdf)

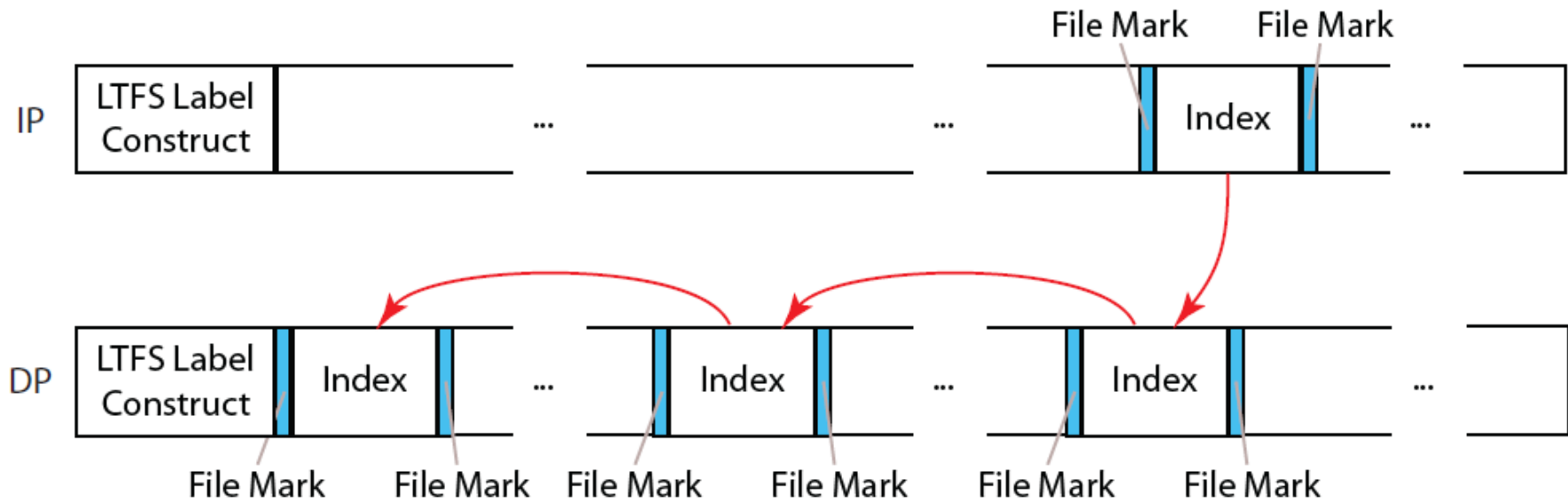
# Sample XML Schema

```
<?xml version="1.0" encoding="UTF-8"?>
<index version="0.9">
  <creator>IBM LTFS 0.20 - Linux - ltfs</creator>
  <volumeuuid>9710d610-5598-442a-8129-48d87824584b</volumeuuid>
  <generationnumber>3</generationnumber>
  <directory>
    <name>LTFS Volume Name</name>
    <creationtime>2010-01-28 19:39:50.715656751 UTC</creationtime>
    <modifytime>2010-01-28 19:39:55.231540960 UTC</modifytime>
    <accesstime>2010-01-28 19:39:50.715656751 UTC</accesstime>
    <contents>
      <directory>
        <name>directory1</name>
        <contents>
          <file>
            <name>binary_file.bin</name>
            <length>10485760</length>
            <extentinfo>
              <extent>
                <partition>b</partition>
                <startblock>8</startblock>
                <byteoffset>0</byteoffset>
                <bytecount>720000</bytecount>
              </extent>
              <extent>
                <partition>b</partition>
                <startblock>18</startblock>
                <byteoffset>0</byteoffset>
                <bytecount>9765760</bytecount>
              </extent>
            </extentinfo>
            <extendedattributes>
              <xattr>
                <key>uservalue</key>
                <value>fred</value>
              </xattr>
            </extendedattributes>
          </file>
          <file>
            <name>read_only_file</name>
            <length>0</length>
            <readonly/>
          </file>
        </contents>
      </directory>
    </contents>
  </directory>
</index>
```

# File Sytem Consistency

- When to re-write Index?
  - Too often:
    - Poor performance
    - Potential for tape wear
  - “Journal” Updates in Data Partition
    - Write interim copies of Index in DP
    - Also creates versioning capability
  - Mount-time options for how often
    - File close (ideal for large files)
    - On timer (better for many small files)
    - Unmount (for those who live dangerously)

# Index Arrangement on LTF5 Tape



# Virtual Extended Attributes

- How to access “unavailable” information?
  - Unsupported dates, uid, start block, volume info, media info
    - Implemented as reserved extended attributes
      - Itfs.backupTime
      - Itfs.fileUID
      - Itfs.startblock...
  - Also allows write-only “commands”
    - e.g. Itfs.sync
  - Full list in LTFS Format Specification

# LTFS in Single Drive Mode

- Shows up like any standard (i.e., disk) file system
  - Directories
  - Files
- Tape contains File/Directory Index in Index Partition
  - XML schema
    - Keeps multiple “generations” (older versions of XML schema)
  - Data files written to Data Partition (usually)
- Small data files can optionally be written to Index Partition
  - Quick access, can be cached at mount time
- Tape index “forgotten” at unmount



# LTFS in Library Mode (LTFS LE)

- Mount Library, not Drive
- LTFS Caches Index of each tape read/written
  - Each volume shows as separate file system folder/directory
  - After mount, all tape directories are viewable, searchable
    - Without mounting any tape
    - LTFS drives automation to mount tape on file read/write
- LTFS can recognize when tape leaves, reenters library
  - Performs consistency check to see if tape index has changed



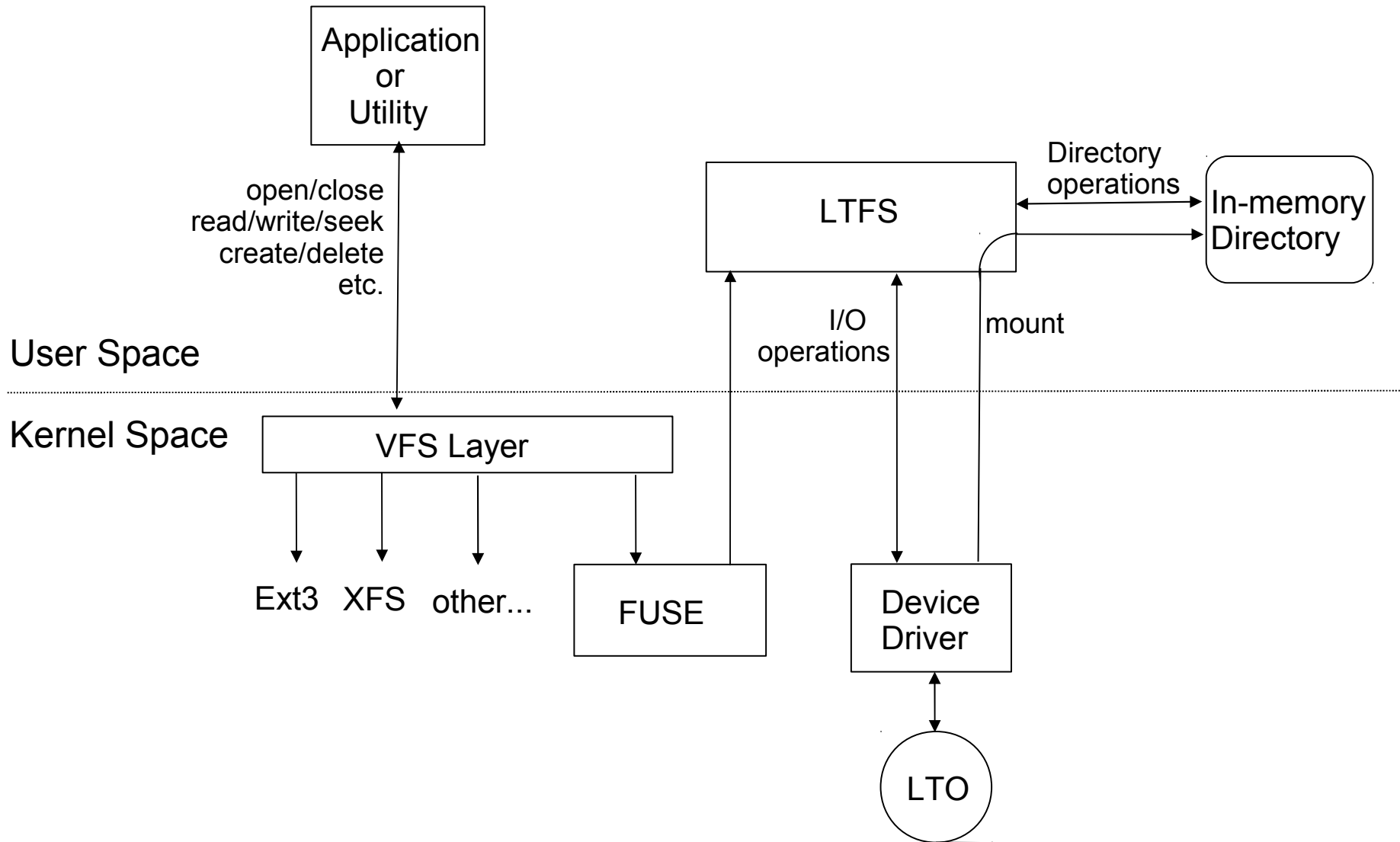
# LTFS in Library Mode (LTFS LE)

- Library mode utility: Itsadmintool
  - Can be used to:
    - vary on/off drives
    - import/export tapes
    - format or check tape volumes
    - get library inventory
  - Command line, scriptable





# LTFS Fuse Implementation



# Thank You!

---

## Questions/Discussion