

Ext3 - What if there is a crash

➔ **Recovery** - Go through log and "redo" operations that have been successfully committed to log

What if ...

- TxBegin but not TxEnd in log?
- TxBegin through TxEnd are in log, but inodes, bitmaps, and data have not yet been checkpointed?
- What if Tx is in log; inodes, bitmaps and data have been checkpointed; but Tx has not been freed from log?

Journaling Modes

Journaling has cost - one write = two disk writes (two seeks with hard disks)

➔ Several journaling modes balance consistency and performance

- **Data journalling** - journal all writes, including file data

- ◉ expensive to journal data

- **Metadata journaling** - journal only metadata

Used by most FS (IBM JFS, SGI XFS, NTFS)

- ◉ file may contain garbage data

- **Ordered mode** - write file data to real FS first, then journal metadata

Default mode for ext3

- ◉ old file may contain new data