Ext3 Support in Mac OS X

Kyle Sun
<skh@cmu.edu>
2009.9.16
Outline

- Problem
- From ext2 to ext3
- Approaches
- My plan
- Questions
Problem

- No ext3 support in Mac OS X.
- There are some good implementations of ext2 in Mac OS X. ext2 in Mac is usable.
- $\text{ext3} = \text{ext2} + \text{journaling}$
From ext2 to ext3

- Ext2 has inconsistency problem when system crashes while writing
- Ext3 = ext2 + journaling
- Ext3 and ext2 is compatible
  - Ext2 driver can mount, r/w ext3 partitions
- What Linux folks did is,
  - copy the whole directory of ext2 and rename it to ext3, replace every 'ext2' to 'ext3' in source code
  - Add journaling
About Journaling

- Journaling is logging
- Makes transaction 'atomic'
- Every op writes to journal first, then writes to data
  1) Write to journal
  2) Commit to journal
  3) Write the content to disk (checkpointing)
Journal format

- Three types of journal blocks
  - Metadata block, entire data block that is modified
  - Descriptor block, describes the metadata blocks.
  - Header block, records current head/tail and SQN. Always the first block in the journal file and fixed location.

- Inode number and device of journal file are stored in ext3 superblock.
More about Journaling

• How it works?

  1) Close transaction, batch ops together
  2) Start flushing transaction to journal via a separated kernel thread
  3) Wait all disk ops to complete as well as all transaction update complete
  4) Commit, update journal header block
  5) Checkpointing, after all blocks in one transaction flushed to disk, update the tail
More about Journaling

• How to recover from crash?
  1) Get the SQN from header block
  2) Scan from the tail to the head, do recovering.
Approaches

- **Kext – Kernel extension**
  - Loadable kernel module in Mac OS X
  - Based on ext2fsx project
  - Full access to the kernel space
  - Hard to debug

- **Fuse – Filesystem in User space**
  - Macfuse and fuse-ext2 project
  - User space, easy to debug
  - Inefficient, limited
My plan

- Kext way

- Several things to do
  - Copy the ext2 folder and rename it to ext3
  - Recover from crashes by reading journals
  - Write to journal, committing & checkpointing
    - Journal layer
  - Utilities: mkfs, fsck, etc
  - License issues
Testing

- Continuously read/write
- Redo after crashes
  Sudden crash while writing large files
  - Linux write and Mac redo
  - Mac write and Linux redo
  - Mac write and Mac redo
- Can do some benchmark test
THANK YOU!