Solaris Lock Manager Changes

Mike Kupfer

kupfer@Eng.Sun.COM

Connectathon 1999

1. Problems with caching/ mmap

- Partial-page locks can cause writes to be lost
  - suppose 2 clients lock and write different bytes of same page

  - client VM wants to write out entire page
  - last one written wins
  - requires protocol fix in general case

Current Approach

- disable caching (and `mmap`) when locked
  - \+ simple
  - – terrible throughput
  - – breaks applications that want to use mmap while file is locked (e.g., nvi)

Possible Fixes

- cache if whole file is locked
  - probably catches most cases, but inflexible
- cache if all locks are page-aligned
  - allows for some concurrent access to same file
- allow locking that doesn't conflict with mapped regions
  - most general, but significantly more bookkeeping

Current Status

- allow caching with page-aligned locks
- prototyped, partially tested
- issue: what if file is mapped and unlock violates page-alignment rule?
  - could fail the unlock, but nobody expects unlocks to fail
  - unmap the file?

2. Cross-Domain Locking and Recovery

- client passes its name in NLM request
- server remembers client's name for callback on reboot
- Solaris hostname has no domain information
- server can't find host name or finds the wrong host (e.g., wanted foo.east, got foo.west)

Possible Fixes

- change Solaris to pass domain information
  - + make Solaris act like everyone else
  - – which domain?
  - – doesn’t deal with legacy clients
- change server to record IP address
  - + handles legacy clients
  - – requires changing X/Open standard protocol (NSM) to tell statd about address
  - – requires changes for, e.g., IPv6

Current Status

- changed server to record IP address
- private sideband protocol to give statd known addresses for a given host name
- statd records host name (file) and network addresses (symbolic links pointing to host name)
- in Solaris 7