FUSE (Filesystem in Userspace) on OpenSolaris (2009/119)

PSARC Inception Review
March 25, 2009
William Krier

FUSE on OpenSolaris (2009/119)

- What is FUSE
- How FUSE Works
- High Level Design Diagram
- FUSE File System API
- FUSE Protocol Specification
- Security Concerns
  - FUSE Mounts
  - Privilege Escalation
  - Block Device Access
  - File System Access Control

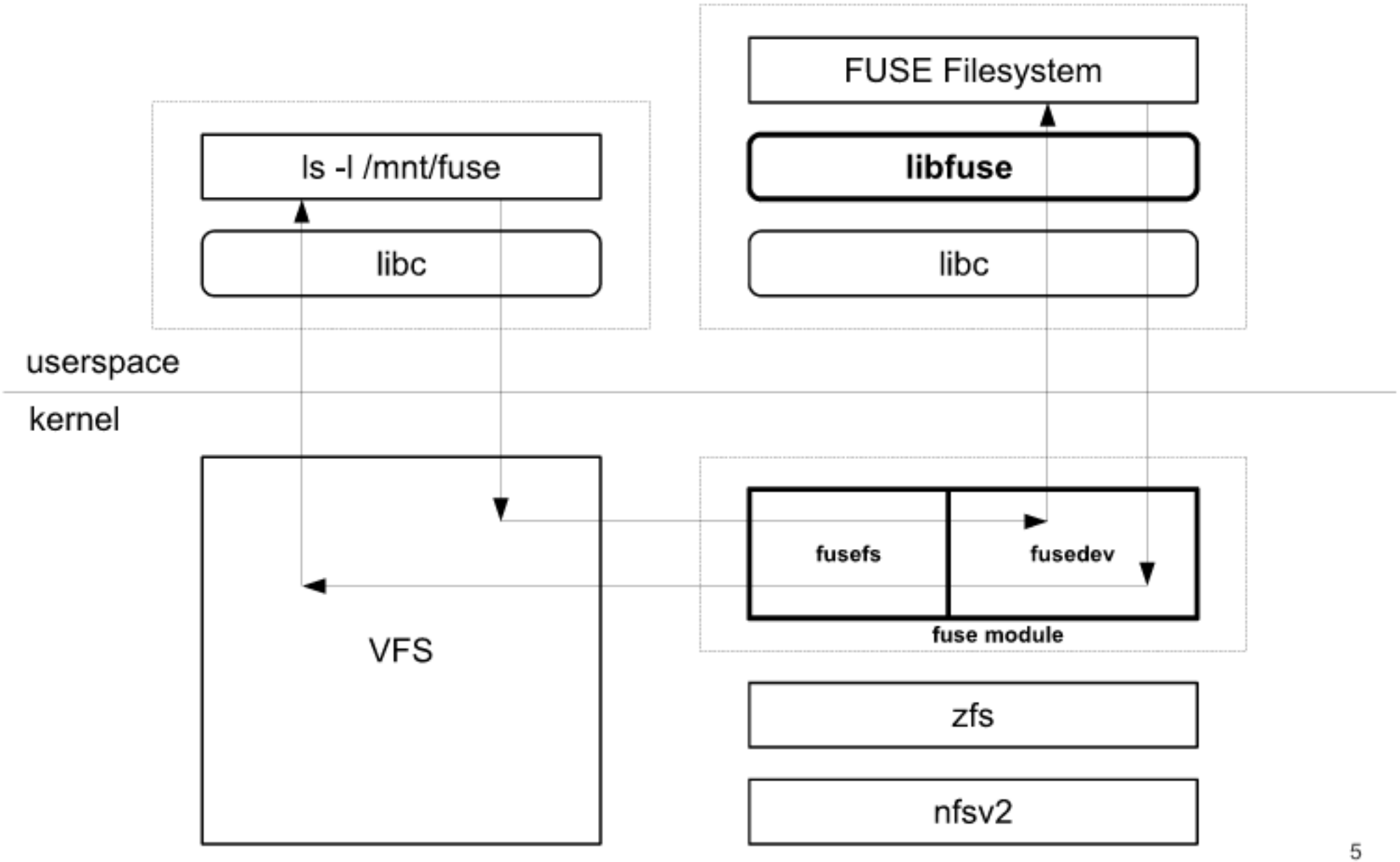
What Is FUSE

- FUSE is a framework which makes it possible to implement a filesystem in a userspace program.
- Features include:
  - simple yet comprehensive API
  - secure mounting by non-root users
  - multi-threaded operation
- Originally developed for Linux and currently runs on 2.4 and 2.6 Linux kernels
- Maintained at fuse.sourceforge.net
- Ported to FreeBSD (fuse4bsd) and Mac OS X
- FUSE library remains the same across platforms

How FUSE Works

- FUSE framework consists of 2 components
  - FUSE user space library (libfuse.so)
    - provides framework and exports FUSE API
  - FUSE kernel module
    - virtual file system (fusefs)
    - character device (/dev/fuse)
- FUSE kernel module redirects vfs calls to the FUSE library via the FUSE character device
- Example FUSE file systems
  - ntfs-3g
  - sshfs
  - davfs

High Level Design Diagram



How FUSE Works (cont.)

- FUSE kernel module registers with VFS.
- FUSE user space file system will link with FUSE library and provides:
  - Register file operation methods w/ library
    - struct fuse_operations
    - getattr, mknod, create, read, write, readdir, readlink, getdir, mknod, chmod, etc.
  - Mount point and options
- FUSE library calls the mount() system call
  - filesystem type is “fuse”
  - filehandle of /dev/fuse passed as option
- Filesystem calls are passed to FUSE library which invokes associated fuse operation in FUSE filesystem.

FUSE File System API

- FUSE file systems use the FUSE API specification to implement necessary file system operations
  - Current version of FUSE API is 2.7
  - Interface is classified as Volatile
  - The API is documented in FUSE_API_Specification.pdf

FUSE Protocol Specification

- Kernel module communicates with FUSE library via the fuse character device
  - During mount, the library opens /dev/fuse and passes file descriptor to kernel.
  - The minor device number is associated with mount point via vfsp->vfs_dev
- Messages are passed via the FUSE device's read/write methods.
- FUSE messages are defined by the FUSE Protocol specification (version 7.8)
  - Protocol is classified as Project Private
  - Protocol is documented in FUSE_Protocol_Specification.pdf

FUSE Security Concerns

- Authorization for non-privileged users to perform file system mounts
- Privilege level escalation for non-privileged mounts
- Access to block devices for non-privileged mounts
- File system-specific access control

FUSE Mounts

- FUSE supports non-privileged mounts
- Accomplished by adding profile to /etc/security/prof_attr
  - FUSE File System Management:::Mount and unmount FUSE filesystems:FUSE.html
- Consist of a single execution authorization in /etc/security/exec_attr
  - granting sys_mount privilege to FUSE mount/unmount programs
  - FUSE File System Management:solaris:cmd:::/usr/bin/fusermount.bin:privs=sys_mount
- Profile must be manually added to users by administrator after installation.

Privilege Escalation

DoS Possibilities with Non-privileged Mounts

- FUSE file systems are only accessible to the user who mounted file system.
  - > Prevents FUSE file system daemon from having ptrace-like capabilities and
  - > Prevents denial of service for the requesting process (stalling system calls)
- Mount option to allow access to other non-root users
- Mount option to allow access to root

Block Device Access

- For FUSE file systems that are backed by block device (ie ntfs-3g)
- Non-root users must have read/write access to block device for non-privilege mount
- Create “fuse” group and add write-allow ACE for specific block device.
- Must be done manually by administrator.

File System Access Control

- By default, FUSE leaves all access control to the file system.
- Mount option to enable FUSE to enforce access control.
  - only allows UNIX-style permission checking, bypassing more sophisticated access controls that may be present in the file system.
  - does not have direct access to file permissions. (only see what file system presents)
- Disposition of file ownership varies by file system.
  - ntfs-3g presents owner as user who mounted fs.
  - sshfs presents owner as reported by remote host.

FUSE on OpenSolaris (2009/119)

William Krier
William.Krier@sun.com